





```

misc_feature      fragment_chain:3"
81047..95004
/note="assembly:fragment:00432
fragment_chain:3"
95105..106663
/note="assembly:fragment:00074
fragment_chain:3"
106764..110896
/note="assembly:fragment:01224
fragment_chain:3"
110597..116311
/note="assembly:fragment:00458
fragment_chain:3"
116412..129147
/note="assembly:fragment:00525
fragment_chain:3"
129248..154765
/note="assembly:fragment:01560
fragment_chain:3"
154866..166501
/note="assembly:fragment:01598
fragment_chain:3"
166602..168873
/note="assembly:fragment:00038
fragment_chain:3"
clone_end:SP6
vector_side:right"
BASE COUNT      45735 a 31934 c 31979 g 53820 t 1405 others
ORIGIN
Query Match      60.6%; Score 235; DB 2; Length 168873;
Best Local Similarity 94.0%; Pred. No. 7.5e-58;
Matches 266; Conservative 0; Mismatches 15; Indels 2; Gaps 2;
QY 1 AACTGCTCAATACGGAACATATTCAGTCCCTCCTCTGCTGCTACAAAGCCGTGATTCCT 1
|||||
DB 26096 AACTGCTCAATATGGAACATATTCAGTCCCTCCTCCTGCTGCTACAAAGCCGTGATTCCT 26157
QY 61 TGCTATGACAGCAAGCTGTGTTAATCTACAGGAGCCATACCTGCTGAACTTTAT 120
|||||
DB 26156 TGTCTATGACAGCAAGCTGTGTTAATCTACAGGAGCCATACCTGCTGAACTTTAT 26215
QY 121 GCTTACAGTGCACACGTGAGTCAGTTCGATTTATATGTTTCAGTCCGTAGCATATTA 180
|||||
DB 26216 GCTTACAGTGCACACGTGAGTCAGTTCGATTTATATGTTTCAGTCCGTAGCATATTA 26275
QY 181 GCATACAGAAACCTTTCATTTCCATTCCTACAGAACTGCACAGGACGATGCTCTACAG 240
|||||
DB 26276 GCT-ACAGAAACCTTTCATTTCCATTCCTACAGAACT-GCACAGGACGATGCTCTACAG 26333
QY 241 GTCTACAAAGAAACCTTCAGATCATCTTCTTGAGGGAGAACAGC 263
|||||
DB 26334 GTCTACAAAGAAACCTTCAGATCATCTTCTTGAGAACAGC 26376
|||||
RESULT 3
AC050848 165555 bp DNA linear chr1 26376-26376
DEFINITION Homo sapiens chromosome 11, clone RP11-76024, complete sequence.
ACCESSION AC050848
VERSION AC050848.4 GI:1795297
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1. (bases 1 to 169555)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-76024
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 169555)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,

```

[illegible]

```

COMMENT
Submitted (28-11-2001) Whitehead Institute/MIT Center for Genome
Research, 220, Whites Street, Cambridge, MA 02141, USA
Note: In this sequence version replaced g1:17367341.
All repeats were identified using RepeatMasker:
http://ftp.cse.washington.edu/KM/RepeatMasker.html
Genes:
Genes: Whitehead Institute/ MIT Center for Genome Research
Genes: COPI, MIR
Web Site: http://www.seq.wi.mit.edu
Contact: s.levine@sumissions.wi.mit.edu
Project Information
Center Project Name: 112557
Center Clon. Name: 760_O_24
FEATURES
Location/Qualifiers
1..112557
/translation "Homo sapiens"
/old_name "Lazou1956c"
/label "c1"
/comp "c1"
/clone "RP11-760v24"
/clone "11b_RP11-11 Human Male BAC"
size 112
note "300 bp SGL region"
765..779
note "300 bp SGL region"

```

repeat_L_region	complement(1350., 1535)
repeat_L_region	/rpt_family="L1HA10"
repeat_L_region	/rpt_family="L1HA4"
repeat_L_region	4020., 4311
repeat_L_region	/rpt_family="L2"
repeat_L_region	complement(446., 5200)
repeat_L_region	/rpt_family="L1MC70"
repeat_L_region	complement(5201., 5472)
repeat_L_region	/rpt_family="A105g"
repeat_L_region	complement(5473., 5545)
repeat_L_region	/rpt_family="L1W70"
repeat_L_region	complement(5700., 5864)
repeat_L_region	/rpt_family="L1W70"
repeat_L_region	5894., 6295
repeat_L_region	/rpt_family="ML11b"
repeat_L_region	complement(6348., 6457)
repeat_L_region	/rpt_family="ML11a1"
repeat_L_region	7251., 7561
repeat_L_region	/rpt_family="L1MEC"
repeat_L_region	7570., 7805
repeat_L_region	/rpt_family="L1MEC"
repeat_L_region	7882., 8200
repeat_L_region	/rpt_family="L1H4"
repeat_L_region	8202., 8766
repeat_L_region	/rpt_family="L1R22"
repeat_L_region	8551., 9677
repeat_L_region	/rpt_family="HERVK22"
repeat_L_region	10047., 10215
repeat_L_region	/rpt_family="HERVK22"
repeat_L_region	10621., 10656
repeat_L_region	/rpt_family="HERVK22"
repeat_L_region	10656., 11735
repeat_L_region	/rpt_family="HERVK22"
repeat_L_region	11805., 12768
repeat_L_region	/rpt_family="HERVK4"
repeat_L_region	12770., 13336
repeat_L_region	/rpt_family="L1R22"
repeat_L_region	13371., 13494
repeat_L_region	/rpt_family="L1H4"
repeat_L_region	13567., 13565
repeat_L_region	/rpt_family="HERS1A"
repeat_L_region	14200., 14617
repeat_L_region	/rpt_family="HERV3"
repeat_L_region	14820., 16078
repeat_L_region	/rpt_family="HERV3"
repeat_L_region	16085., 18592
repeat_L_region	/rpt_family="HERV3"
repeat_L_region	18592., 19341
repeat_L_region	/rpt_family="HERV3"
repeat_L_region	19622., 19673
repeat_L_region	/rpt_family="L1R4"
repeat_L_region	19763., 20051
repeat_L_region	/rpt_family="L1R4"
repeat_L_region	20136., 20633
repeat_L_region	/rpt_family="MERS1A"
repeat_L_region	20813., 20878
repeat_L_region	/rpt_family="M1R"
repeat_L_region	complement(21650., 21798)
repeat_L_region	/rpt_family="HER5A"
repeat_L_region	22492., 22591
repeat_L_region	/rpt_family="L1R35"
repeat_L_region	22714., 23046
repeat_L_region	/rpt_family="L1R35"
repeat_L_region	23456., 23816
repeat_L_region	/rpt_family="L2"
repeat_L_region	complement(23801., 23858)
repeat_L_region	/rpt_family="L2"
repeat_L_region	24476., 24582
repeat_L_region	/rpt_family="L2"
repeat_L_region	25186., 25278
repeat_L_region	/rpt_family="L1ME1"
repeat_L_region	25309., 26522

[illegible][illegible]







[illegible]





QY 361 TCCCTGATCCATTAANGTAT 382  
DB 61021 DEGVSYSWTDHMSKMBRD 61042

## RESULT 11

AP000576/C

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

121922 bp DNA contig of Homo sapiens chromosome 11 clone CH9-6J10 map 11.12. WORKING DRAFT.

AP000576.4 GI:9927276  
HTG: HTGS\_PHASE1: HTGS\_DRAFT.  
Homo sapiens DNA, clone:CH9-6J10.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 121922)  
Hattori, M., Ishii, K., Toyoda, A., Taylor, F. B., Hong, S.-E., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H., and Sakaki, Y. Homo sapiens 121,922 genomic DNA of 11q12  
Published Only in Database (1999)  
2 (bases 1 to 121922)  
Hattori, M., Ishii, K., Toyoda, A., Taylor, F. B., Hong, S.-E., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H., and Sakaki, Y. Direct Submission  
Submitted (08-Oct-1999) Masahiro Hattori, The Institute of Chemical and Chemical Research (RIKEN), Genomic Sciences Center (GSC) 1-15-1 Kitasato, Sayamihara, Kanagawa 248-0555, Japan  
(E-mail: hattori@qsc.riken.go.jp, URL: http://hap.qsc.riken.go.jp, Tel: 81-42-778-9923, Fax: 81-42-778-9924)  
On Aug 26, 2000 this sequence version replaced gi:119274.

Center: RIKEN Genomic Sciences Center (GSC)  
Center code: RIKEN  
Web site: http://hap.qsc.riken.go.jp/  
Contact: hattori@qsc.riken.go.jp

Project Information  
Center project name: HumDraft1  
Center clone name: CH9-6J10

Summary Statistics  
Sequencing vector: PCR products; 1067 of reads  
Chemistry: Dye-terminator ET-amersham; 1008 of reads  
Assembly program: Phrap; version 0.950319  
Consensus quality: 118280 bases at least 40  
Consensus quality: 119669 bases at least 40  
Consensus quality: 120421 bases at least 40  
Insert size: 120822; sum-of-contigs  
Quality coverage: 12.41x in 920 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 25319 contig of 25319 bp in length  
25420 43536 contig of 18117 bp in length  
43537 62163 contig of 18527 bp in length  
62164 74470 contig of 12207 bp in length  
74471 84826 contig of 10256 bp in length  
84827 94411 contig of 9485 bp in length  
94412 103560 contig of 9049 bp in length  
103561 111843 contig of 8183 bp in length  
111844 116839 contig of 4896 bp in length  
116840 119647 contig of 2708 bp in length  
119648 120811 contig of 1064 bp in length  
120812 121922 contig of 1011 bp in length.

\* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces is not known and their order in this sequence record is

\* arbitrarily. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 25319 contig of 25319 bp in length  
25420 43536 contig of 18117 bp in length  
43537 62163 contig of 18527 bp in length  
62164 74470 contig of 12207 bp in length  
74471 84826 contig of 10256 bp in length  
84827 94411 contig of 9485 bp in length  
94412 103560 contig of 9049 bp in length  
103561 111843 contig of 8183 bp in length  
111844 116839 contig of 4896 bp in length  
116840 119647 contig of 2708 bp in length  
119648 120811 contig of 1064 bp in length  
120812 121922 contig of 1011 bp in length.

Location/altalters  
1. 121922

/organism: "Homo sapiens"  
/db\_xref: "taxon:9606"

/chromosome: "11"  
/map: "11q12"

/clone: "CH9-6J10"

1. 25319  
/note: "assembly-fragment"

25420. 43536  
/note: "assembly-fragment"

43537. 62163  
/note: "assembly-fragment"

62164. 74470  
/note: "assembly-fragment"

74471. 84826  
/note: "assembly-fragment"

84827. 94411  
/note: "assembly-fragment"

94412. 103560  
/note: "assembly-fragment"

103561. 111843  
/note: "assembly-fragment"

111844. 116839  
/note: "assembly-fragment"

116840. 119647  
/note: "assembly-fragment"

119648. 120811  
/note: "assembly-fragment"

120812. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

121922. 121922  
/note: "assembly-fragment"

```

RESULT 12
AP001636 161582 bp DNA linear F01:05-12-01-01
LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-142C4
DEFINITION complete sequence.
ACCESSION AP001636
VERSION AP001636.4 GI:17425232
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:RP11-142C4.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2000)
AUTHORS 2 (bases 1 to 161582)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Shuhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 226-0045, Japan
(E-mail:hattori@psc.riken.go.jp, URL:http://wbp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Dec 7, 2001 this sequence version replaced gi:15927250.

COMMENT
FEATURES
source
location/Qualifiers
1..161582
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-142C4"
BASE COUNT 48492 a 31636 c 32205 g 49245 t
ORIGIN
Query Match 11.6% Score 45.2; DB 9; Length 161582;
Best Local Similarity 68.8%; Pred No. 0.022;
Matches 77; Conservative 0; Mismatches 33; Indels -; Gaps 1;

Cy 255 TTCAGATCATCTCTTGAGGAAGAG-CTGAAGTCTCATATAGATGCTTTCTTC 31-
11111 1111111 1 11 11111 11 111 11 11
Db 46166 TTCAGAGTTCTTCTTGATGCTGCTCAAGCATCCCAAAATGATGATGATG 41-13
11111 1111111 11111 11 1111 11 1111 11
Cy 313 ATAAGCTCAGAGCTGCAGATTCGTATAATCTTAAAGAGGATGCG 34-
111111 11111111 11111 11 1111 11 1111 11
Db 46226 ATAAGCTCCAGAGCTGCTGTTCTATACAAACCTTGGCTCGACATGAT 42-77
111111 11111111 11111 11 1111 11 1111 11

RESULT 13
AC090309 161586 bp DNA linear F01:14-20-01-01
LOCUS Homo sapiens chromosome 11, clone RP11-142C4, complete sequence.
DEFINITION AC090309
ACCESSION AC090309
VERSION AC090309.4 GI:15055316
KEYWORDS HTG.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavsky,L., Bouckhalter,B., Brown,A.,
Camarata,J., Campopiano,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S.,
Collumore,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,T., Galagan,J.,
Gallardo,S., Gault,S., Goffette,M., Graham,L., Grand-Pierre,N.,
Hais,B., Heideck,H., Hulton,L., Hulme,W., Iller,I., Johnson,R.,
Jones,C., Katsas,A., Lacroque,K., Lamazares,R., Landers,T.,
Lewicki,J., Levine,K., Liu,G., Maclean,C., Macdonald,P.,
Margolis,N., Matthews,C., McCarthy,M., McEwan,P., McKenna,K.,
McPherson,K., Melrim,J., Menus,L., Mihova,I., Miya,Y.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Reilly,D., Oliver,J., Peterson,K.,
Plunkhant,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rietack,M., Riley,K., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
Senguez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Tronis,N., Trigglio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,H.J., Young,G., Zalcov,J.,
Zemke,L., Zimmer,A. and Zou,Y.M.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 32 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 161586)
Barna,N., Bastien,V., Blom,T., Boguslavsky,L., Bouckhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collumore,A.,
Cooke,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,T., Galagan,J.,
Gallardo,S., Gault,S., Graham,L., Grand-Pierre,N., Hayes,B.,
Hertel,L., Hulme,W., Iller,I., Johnson,R., Jones,C., Kamal,A.,
Karatas,A., Kells,L., Landers,T., Levine,K., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Melrim,J., Menus,L., Mihova,I., Miya,Y.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,K., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Reilly,D., Oliver,J., Peterson,K.,
Plunkhant,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zalcov,J.,
Zemke,L., Zimmer,A. and Zou,Y.M.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2002) Whitehead Institute/MIT Center for Genome
Research, 32 Charles Street, Cambridge, MA 02141, USA
On Aug 1, 2001 this sequence version replaced gi:1159564.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Smit,A.F.A. & Green,P. (1996-1997)
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center code name: 142C4
Center clone name: 142C4

FEATURES
source
location/Qualifiers
1..161586
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-142C4"
/clone_lib="RP11-11 Human Male BAC"
/complement(405..596)
/rpl_family="L2"
/complement(85..300)
/rpl_family="L193e"
/ncbi="1631
/ncbi="30 qual SHL region"
/ncbi="1755..1759
/ncbi="330 qual SHL region"
/ncbi="3177..3326
/rpl_family="L194u"
/ncbi="3597..3637

```

```

REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavsky,L., Bouckhalter,B., Brown,A.,
Camarata,J., Campopiano,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S.,
Collumore,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,T., Galagan,J.,
Gallardo,S., Gault,S., Goffette,M., Graham,L., Grand-Pierre,N.,
Hais,B., Heideck,H., Hulton,L., Hulme,W., Iller,I., Johnson,R.,
Jones,C., Katsas,A., Lacroque,K., Lamazares,R., Landers,T.,
Lewicki,J., Levine,K., Liu,G., Maclean,C., Macdonald,P.,
Margolis,N., Matthews,C., McCarthy,M., McEwan,P., McKenna,K.,
McPherson,K., Melrim,J., Menus,L., Mihova,I., Miya,Y.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Reilly,D., Oliver,J., Peterson,K.,
Plunkhant,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rietack,M., Riley,K., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
Senguez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Tronis,N., Trigglio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,H.J., Young,G., Zalcov,J.,
Zemke,L., Zimmer,A. and Zou,Y.M.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 32 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 161586)
Barna,N., Bastien,V., Blom,T., Boguslavsky,L., Bouckhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collumore,A.,
Cooke,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,T., Galagan,J.,
Gallardo,S., Gault,S., Graham,L., Grand-Pierre,N., Hayes,B.,
Hertel,L., Hulme,W., Iller,I., Johnson,R., Jones,C., Kamal,A.,
Karatas,A., Kells,L., Landers,T., Levine,K., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Melrim,J., Menus,L., Mihova,I., Miya,Y.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,K., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Reilly,D., Oliver,J., Peterson,K.,
Plunkhant,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zalcov,J.,
Zemke,L., Zimmer,A. and Zou,Y.M.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2002) Whitehead Institute/MIT Center for Genome
Research, 32 Charles Street, Cambridge, MA 02141, USA
On Aug 1, 2001 this sequence version replaced gi:1159564.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Smit,A.F.A. & Green,P. (1996-1997)
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center code name: 142C4
Center clone name: 142C4

FEATURES
source
location/Qualifiers
1..161586
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-142C4"
/clone_lib="RP11-11 Human Male BAC"
/complement(405..596)
/rpl_family="L2"
/complement(85..300)
/rpl_family="L193e"
/ncbi="1631
/ncbi="30 qual SHL region"
/ncbi="1755..1759
/ncbi="330 qual SHL region"
/ncbi="3177..3326
/rpl_family="L194u"
/ncbi="3597..3637

```

[illegible]



BASE COUNT 472 a 336 c 394 g 140 t 5 others

ORIGIN

Query Match  
Best Local Similarity 11.1% Score 43 DB 6 Length 1627;  
Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 283 CTGAAGCTACATGATCGTGTGGCTTCATACTCAGAACGUSAGAATCTGATAA 342  
||||| 111 111 11 111111 111111 111111  
Db 310 CTGAAGCATCCCAAGAAATGCATCTAGCAATAAATCCCATTAAGCTGGCGCTATACAA 369  
||||| 111 111 111111 111111 111111 111111  
QY 343 AATCGCTTAGAAAAGACATCCCTCATCTCAATCCATAAANGTATATGCC 367  
||||| 111111 111111 111111 111111 111111 111111  
Db 370 AATCGCTGGCCAGAGCAT-CCCTGATGTCCTCGAAGTGCTATGCC 413

RESULT 15  
BC008353 1646 bp mRNA linear PRI 12-JUL-2001  
LOCUS Homo sapiens, similar to RIKEN cDNA O61008P16 gene, clone  
MGCL15937 IMAGE:3537224, mRNA, complete cds.  
ACCESSION BC008353  
VERSION BC008353.1 GI:14249941  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 1646)  
Straussberg,R.  
Direct Submission  
Submitted (23-MAY-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2530,  
USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [gcgpb@ncl.nih.gov](mailto:gcgpb@ncl.nih.gov)  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcsc.bc.ca](mailto:info@bcsc.bc.ca)  
Susanna Jones, Jennifer Asano, Ian Bosdet, Yaron Buterfield,  
Suzanna Chan, Readman Chin, Chris Pjelli, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Keta Kusche, Oliver Lee, Sco  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLevy, Steven  
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saadati, Jacqueline  
Schell, Duane Smalls, Michael Smith, Lorraine Spencer, Jeff Stolt,  
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

REMARK COMMENT

FEATURES  
Source  
1..1646  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:15937 IMAGE:3537224"  
/tissue-type="lung small cell carcinoma"

```

/clone_110="H1.M2.7"
/AltHost="E8105-R"
/note="Vector: pGB7"
341..1249
/contig_start=1
/product="Similar to RIKEN cDNA 661008p16 gene"
/protein_id="AA08353.1"
/db_xref="GI:14249542"
/translation="MILINNSHLKILATKSLRSPESLKYGSVYHINGNPNPMNV
LVDNPEYQGIKPKQEMIDMSYTNVRMSKEPQSEVLKCEIVNNKORLO
LGDGSESGEIKVAIPESKSVKVESRLLIVTDILEINSSSKIGSMAPGPPD
EESSETPFKYADLDVYSGLVNNKRGRESLHYTKETDLPACMLGPGGVY
SNVIMDPSCEYGMAYSMKRYKRTGMAVNVKRYKYLKQKNIPYISVLENEEDSRKF
VGDVDFEASGEEMVGYCTPQNTLPF"
BASE COUNT      506 a      329 c      389 g      422 t
ORIGIN
Query Map: 111% Score 43; DB 9; Length 1646;
Best Local Similarity 56.6%; Pred. No. 0.093;
Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;
Q# 283 11GAGAGTGCACATAAAS13CTTGTGCTGATACCTGCAGAGCTGCAGATCTGATA 342
111111 11 111 1 11 111111 11 111111 11 11111 1 11 11 1
D# 324 CTGAGGACATCCACACAAATAGATCTACTGATGATAAATCCCAATAGS13CTGACCCTATACA 383
CTGAGGACATCCACACAAATAGATCTACTGATGATAAATCCCAATAGS13CTGACCCTATACA 383
Q# 313 ATACCTTGAAGAAAGAGCATCCCTCCCTGATCATCATCATATATATATGCG 367
111111 1 111111 111111 111 11 11 11 1
D# 384 AATCCTTGCCTAGAGAAAT-CCCTGAGTCCCTCGAAGGTGATATGCG 427
AATCCTTGCCTAGAGAAAT-CCCTGAGTCCCTCGAAGGTGATATGCG 427

```

Search completed: January 2, 2003, 15:10:05  
JCL time : 2650 secs



XX Claim 6; Page 39; 45pp; English.

XX This sequence represents a breast specific gene (BSG) clone. The bsg  
XX sequences are also referred to as breast specific markers (BSM). The  
XX invention relates to methods for diagnosing, monitoring, staging, treating  
XX and treating breast cancer (BC). The methods comprise measuring the  
XX levels of BSG products in cells, tissues or body fluids of the patient  
XX and comparing the measured levels of BSG with BSG levels of a normal  
XX human control. An antibody against the BSG sequences can be labelled and  
XX used for imaging BC in a patient. The antibody can be conjugated to a  
XX cytotoxic agent, and used for treating BC in a patient.

XX Sequence 388 BP; 115 A; 83 C; 77 G; 111 T; 2 other;

XX Query Match 99.5%; Score 386; DB 21; Length 388;  
XX Best Local Similarity 100.0%; Pred. No. 6,6e-110; Gaps 0;  
XX Matches 386; Conservative 0; Mismatches 0; Indels 0;

1 AGCTGCTCAATACGACATATTCCTACGTCCTGCTGCTACAAACCTGTGATTCCT 4  
1 AAGTCTCAATACGACATATTCCTACGTCCTGCTGCTACAAACCTGTGATTCCT 4  
61 TGCTATGACAGAGACGCTGCTGTTATCTACAGAGACCAACTTCTGAGGCTTAT 120  
61 TGCTATGACAGAGACGCTGCTGTTATCTACAGAGACCAACTTCTGAGGCTTAT 120  
121 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 180  
121 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 180  
181 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 180  
181 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 180  
241 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 240  
241 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 240  
301 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 360  
301 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 360  
361 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 360  
361 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 360

RESULT 2

AAH76198 standard; cDNA; 1645 BP.

AAH76198;

29-OCT-2001 (first entry)

Human drug metabolizing enzyme encoding cDNA (ID No. 388161.3).

Drug metabolizing enzymes; DME; immunosuppressive; cytostatic; antitumor;  
hepatocellular; antileukemic; antitubercular; antibacterial; antiviral;  
antitense therapy; gene therapy; human; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 391..1275

W0200159127-A2.

16-AUG-2001.

08-FEB-2001; 2001WO-0504423.

XX 13-FEB-2001; 2001WO-0504423;  
XX 25-FEB-2001; 2001WO-0504423;  
XX 03-MAR-2001; 2001WO-0504423;  
XX 09-MAR-2001; 2001WO-0504423;  
XX 17-MAR-2001; 2001WO-0504423;  
XX (100%) cDNA; 388161.3 (100%)  
XX 13-FEB-2001; 2001WO-0504423;  
XX 25-FEB-2001; 2001WO-0504423;  
XX 03-MAR-2001; 2001WO-0504423;  
XX 09-MAR-2001; 2001WO-0504423;  
XX 17-MAR-2001; 2001WO-0504423;  
XX (100%) cDNA; 388161.3 (100%)  
XX 13-FEB-2001; 2001WO-0504423;  
XX 25-FEB-2001; 2001WO-0504423;  
XX 03-MAR-2001; 2001WO-0504423;  
XX 09-MAR-2001; 2001WO-0504423;  
XX 17-MAR-2001; 2001WO-0504423;  
XX (100%) cDNA; 388161.3 (100%)

XX This sequence represents a breast specific gene (BSG) clone. The bsg  
XX sequences are also referred to as breast specific markers (BSM). The  
XX invention relates to methods for diagnosing, monitoring, staging, treating  
XX and treating breast cancer (BC). The methods comprise measuring the  
XX levels of BSG products in cells, tissues or body fluids of the patient  
XX and comparing the measured levels of BSG with BSG levels of a normal  
XX human control. An antibody against the BSG sequences can be labelled and  
XX used for imaging BC in a patient. The antibody can be conjugated to a  
XX cytotoxic agent, and used for treating BC in a patient.

XX Sequence 388 BP; 115 A; 83 C; 77 G; 111 T; 2 other;

XX Query Match 99.5%; Score 386; DB 21; Length 388;  
XX Best Local Similarity 100.0%; Pred. No. 6,6e-110; Gaps 0;  
XX Matches 386; Conservative 0; Mismatches 0; Indels 0;

1 AGCTGCTCAATACGACATATTCCTACGTCCTGCTGCTACAAACCTGTGATTCCT 4  
1 AAGTCTCAATACGACATATTCCTACGTCCTGCTGCTACAAACCTGTGATTCCT 4  
61 TGCTATGACAGAGACGCTGCTGTTATCTACAGAGACCAACTTCTGAGGCTTAT 120  
61 TGCTATGACAGAGACGCTGCTGTTATCTACAGAGACCAACTTCTGAGGCTTAT 120  
121 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 180  
121 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 180  
181 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 180  
181 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 180  
241 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 240  
241 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 240  
301 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 360  
301 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 360  
361 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 360  
361 GCTTACAGTACAGACGCTGAGTCACTGTAATTTATGCTGTTGAGTCAAGATATA 360

Human drug metabolizing enzyme encoding cDNA (ID No. 388161.3).

Drug metabolizing enzymes; DME; immunosuppressive; cytostatic; antitumor;  
hepatocellular; antileukemic; antitubercular; antibacterial; antiviral;  
antitense therapy; gene therapy; human; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 391..1275

W0200159127-A2.

16-AUG-2001.

08-FEB-2001; 2001WO-0504423.

```

XX DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:15.
XX
XX KW Human: prostate cancer; prostate cancer antigen; detection; diagnosis;
XX KW neuroprotective; cytostatic; cardioactive; immunomodulatory;
XX KW anti-infective; anti-inflammatory; nephroprotective;
XX KW anti-bacterial; gene therapy; neutral; immune; reproductive; renal;
XX KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX KW wound; infectious disease; SS.
XX
XX CS Homo sapiens.
XX
XX PN WO20005174-A1.
XX
XX PD 21-SEP-2000.
XX
XX PF 08-MAR-2000; 2000MO-US05988.
XX
XX PR 12-MAR-1999; 99US-0124270.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX
XX PI Rosen CA, Ruben SM;
XX
XX DR WPI; 2000-567513/55.
XX LR P-PSDB; AAB56487.
XX
XX PT Prostate cancer associated gene sequences, referred to as prostate
XX PT cancer antigens, useful for treatment, prevention, and diagnosis of
XX PT disorders such as prostate cancer.
XX
XX PS Claim 1: Page 693-694; 2338pp; English.
XX
XX CC AAF15566 to AAF16505 encode the human prostate cancer associated
XX CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX CC The prostate cancer antigens can have neuroprotective, cytostatic,
XX CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX CC nephroprotective, anti-infective, gynaecological and antibacterial activities.
XX CC and can be used in gene therapy. The prostate cancer antigen
XX CC polynucleotides may be used for detection of prostate cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The prostate cancer antigens may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX CC AAB57303 represent sequences used in the exemplification of the present
XX CC invention.
XX
XX SQ Sequence 1189 BP; 333 A; 245 C; 300 G; 306 T; 3 other:
XX
XX Query Match 11.1%; Score 43; DB 21; Length 1189;
XX Best Local Similarity 68.6%; Pred. No. 0.0039;
XX Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 0
XX
XX CY 263 CTGAAGTCTCATATAGATGCTTGTGCTCACTACTCAAGAGCTGAGATATATATA 442
XX 111111 111111 111111 111111 111111 111111
XX DB 303 CTGAAGCATCCACAGATGATCTACTGAAATATCCCAATAGCTGATGCTCACTACA 442
XX 111111 111111 111111 111111 111111 111111
XX CY 343 AATCCTTAGAAGAGAGCATCCGATGATCCATTAAGATATATATATG 367
XX 111111 111111 111111 111111 111111 111111
XX DB 363 AATCCTTGCCAGAGAGCAT-CCCTTAGTCCCTGGAAGGTGATATGCT 406
XX
XX RESULT 4
XX ID AAV24501 standard: cDNA; 1390 BP.
XX AC AAV24501:
XX
XX DT 16-SEP-2002 (first entry)
XX
XX DE Human prostate expression marker cDNA 24492.

```

```

XX KW Human: prostate cancer; prostate cancer antigen; detection; diagnosis;
XX KW neuroprotective; cytostatic; cardioactive; immunomodulatory;
XX KW anti-infective; anti-inflammatory; nephroprotective;
XX KW anti-bacterial; gene therapy; neutral; immune; reproductive; renal;
XX KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX KW wound; infectious disease; SS.
XX
XX CS Homo sapiens.
XX
XX PN WO20005174-A1.
XX
XX PD 21-SEP-2000.
XX
XX PF 08-MAR-2000; 2000MO-US05988.
XX
XX PR 12-MAR-1999; 99US-0124270.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX
XX PI Rosen CA, Ruben SM;
XX
XX DR WPI; 2000-567513/55.
XX LR P-PSDB; AAB56487.
XX
XX PT Prostate cancer associated gene sequences, referred to as prostate
XX PT cancer antigens, useful for treatment, prevention, and diagnosis of
XX PT disorders such as prostate cancer.
XX
XX PS Claim 1: Page 693-694; 2338pp; English.
XX
XX CC AAF15566 to AAF16505 encode the human prostate cancer associated
XX CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX CC The prostate cancer antigens can have neuroprotective, cytostatic,
XX CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX CC nephroprotective, anti-infective, gynaecological and antibacterial activities.
XX CC and can be used in gene therapy. The prostate cancer antigen
XX CC polynucleotides may be used for detection of prostate cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The prostate cancer antigens may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX CC AAB57303 represent sequences used in the exemplification of the present
XX CC invention.
XX
XX SQ Sequence 1189 BP; 333 A; 245 C; 300 G; 306 T; 3 other:
XX
XX Query Match 11.1%; Score 43; DB 21; Length 1189;
XX Best Local Similarity 68.6%; Pred. No. 0.0039;
XX Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 0
XX
XX CY 263 CTGAAGTCTCATATAGATGCTTGTGCTCACTACTCAAGAGCTGAGATATATATA 442
XX 111111 111111 111111 111111 111111 111111
XX DB 303 CTGAAGCATCCACAGATGATCTACTGAAATATCCCAATAGCTGATGCTCACTACA 442
XX 111111 111111 111111 111111 111111 111111
XX CY 343 AATCCTTAGAAGAGAGCATCCGATGATCCATTAAGATATATATATG 367
XX 111111 111111 111111 111111 111111 111111
XX DB 363 AATCCTTGCCAGAGAGCAT-CCCTTAGTCCCTGGAAGGTGATATGCT 406
XX
XX RESULT 5
XX ID AAV24501 standard: cDNA; 1390 BP.
XX AC AAV24501:
XX
XX DT 16-SEP-2002 (first entry)
XX
XX DE Human prostate expression marker cDNA 24492.

```

```

AM pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
XX
XX 16-MAR-2000; 2000US-189862P.
XX
XX 25-MAY-2000; 2000US-207454P.
XX
XX 09-JUN-2000; 2000US-211314P.
XX
XX 18-JUL-2000; 2000US-215007P.
XX
XX 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662755/76.
XX
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX
XX Claim 1; Page 4654; 11750pp; English.
XX
XX
XX The invention relates to an isolated nucleic acid molecule (i) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV52213) of the
XX specification or its complement. (1) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX
XX Sequence 1390 BP; 384 A; 328 C; 350 G; 321 T; 7 other:
XX
XX
XX Query Match 11.1%; Score 43; DB 23; Length 1390;
XX Best Local Similarity 68.6%; Pred. No. 0.0042;
XX Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;
XX
XX 283 CTGAGTGCATCAATGAGTCTGCTTCAATACCTCAGAGAGCTGAGATTCGTATTA 342
XX ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 362 CTGAGATCCACAGAGATGATCTACTGATTAATCAATCAATCAATCAATCAATCA 421
XX ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 343 AATCTTAGAAGAGAGATCCCTGATTCATTAANGATATATTCG 387
XX ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 422 AATCTTGCCAGAGAGCAT-CCCTGAGTCCCTGAGAGGATATATTCG 465
XX
XX
XX RESULT 6
XX ABV24715
XX ID ABV24715 standard; cDNA; 1390 BP.
XX
XX
XX ABV24715;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 24706.
XX
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX

```

```

S Homo sapiens.
XX
XX W-0160860-A2.
XX
XX 23-AUG-2001.
XX
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
XX
XX 16-MAR-2000; 2000US-189862P.
XX
XX 25-MAY-2000; 2000US-207454P.
XX
XX 09-JUN-2000; 2000US-211314P.
XX
XX 18-JUL-2000; 2000US-215007P.
XX
XX 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege W, Monahan JE;
XX
XX WPI; 2001-662755/76.
XX
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX
XX Claim 1; Page 4713; 11750pp; English.
XX
XX
XX The invention relates to an isolated nucleic acid molecule (i) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV52213) of the
XX specification or its complement. (1) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX
XX Sequence 1390 BP; 384 A; 328 C; 350 G; 321 T; 7 other:
XX
XX
XX Query Match 11.1%; Score 43; DB 23; Length 1390;
XX Best Local Similarity 68.6%; Pred. No. 0.0042;
XX Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;
XX
XX 283 CTGAGTGCATCAATGAGTCTGCTTCAATACCTCAGAGAGCTGAGATTCGTATTA 342
XX ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 362 CTGAGATCCACAGAGATGATCTACTGATTAATCAATCAATCAATCAATCAATCA 421
XX ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 343 AATCTTAGAAGAGAGATCCCTGATTCATTAANGATATATTCG 387
XX ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 422 AATCTTGCCAGAGAGCAT-CCCTGAGTCCCTGAGAGGATATATTCG 465
XX
XX
XX RESULT 6
XX ABV2539
XX ID ABV2539 standard; cDNA; 1390 BP.
XX
XX
XX ABV2539;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 25372.
XX
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX
XX Homo sapiens.
XX

```

PH WO200160660-A2.  
XX  
PD 23-AUG-2001.  
XX  
XX 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183119P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
FA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI: 2001-662795/76.  
XX  
PI Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1: Page 5013; 11750pp: English.  
XX  
PS The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (1) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;  
CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SU Sequence 1350 BP; 364 A; 328 C; 350 G; 321 T; 7 other:  
XX  
Query Match 11.1%; Score 43; DB 23; Length 1350;  
Best Local Similarity 66.6%; Pred. No. 0.0042;  
Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;  
XX  
QY 283 CTGAGTCTGCATAGATGCTTGTCTTCATAGCTCTGAGAGCTGCAATTCCTGATA 342  
DB 362 CTGAGCATCCACAGACATGATCTTACTGATTAATCTCCATAGCTCTGAGCATATACA 421  
XX  
QY 343 AATCCTTAGAAAAGACATCCCTGATCCATTAAGTATATGCG 387  
DB 422 AATCCTTGCCAGAGACAT-CCCTGAGTCCCTGAGAGTGTATAGCG 465  
XX  
RESULT 8  
ABV27695  
ID ABV27695 standard; CDNA: 1350 BP.  
XX  
AC ABV27695;  
XX  
XX 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker CDNA 27686.  
XX  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW Pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WU200160660-A2.  
XX

PH WO200160660-A2.  
XX  
PD 23-AUG-2001.  
XX  
XX 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183119P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
FA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI: 2001-662795/76.  
XX  
PI Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1: Page 5013; 11750pp: English.  
XX  
PS The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (1) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;  
CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SU Sequence 1350 BP; 364 A; 328 C; 350 G; 321 T; 7 other:  
XX  
Query Match 11.1%; Score 43; DB 23; Length 1350;  
Best Local Similarity 66.6%; Pred. No. 0.0042;  
Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;  
XX  
QY 283 CTGAGTCTGCATAGATGCTTGTCTTCATAGCTCTGAGAGCTGCAATTCCTGATA 342  
DB 362 CTGAGCATCCACAGACATGATCTTACTGATTAATCTCCATAGCTCTGAGCATATACA 421  
XX  
QY 343 AATCCTTAGAAAAGACATCCCTGATCCATTAAGTATATGCG 387  
DB 422 AATCCTTGCCAGAGACAT-CCCTGAGTCCCTGAGAGTGTATAGCG 465  
XX  
RESULT 9  
ABV27695  
ID ABV27695 standard; CDNA: 1627 BP.  
XX  
AC ABV27695;  
XX  
XX 28-JUN-2002 (first entry)  
XX  
DE Novel human coding sequence. Size 1627 bp.  
XX  
XX Human; anti-infective; cytostatic; anti-inflammatory; immunomodulator;  
XX  
XX anti-infective; cytostatic; anti-inflammatory; immunomodulator;  
KW neoplastic; anti-infective; cytostatic; anti-inflammatory; immunomodulator; expressed sequence tag; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WU200160660-A2.  
XX



```

XX AC ABV42132:
XX XX
XX DE 16-SEP-2002 (first entry)
XX DE Human prostate expression marker CDNA 42123.
XX KN Human; prostate cancer; cytostatic; carcinogen; pharmacofarman; marker;
XX KN pharmacogenomic marker; gene; ss.
XX CS Homo sapiens.
XX FN WO200106060-A2.
XX PD 23-AUG-2001.
XX PE 20-FEB-2001; 2001WO-US05171.
XX FR 17-FEB-2000; 2000US-183319P.
XX FR 16-MAR-2000; 2000US-185862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 05-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege NO, Monahan JE;
XX DR WPI: 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PS prostate cells and correlating with presence of prostate cancer, useful
XX PS for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 8445; 11750pp; English.
XX XX
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV6213) of the
XX CC specification or its complement, (I) is useful for:
XX CC (a) assessing whether a patient is afflicted with prostate cancer;
XX CC (b) monitoring the progression of prostate cancer in a patient;
XX CC (c) assessing the efficacy of a test compound to inhibit prostate
XX CC cancer in a patient;
XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX CC in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound;
XX CC (g) determining whether prostate cancer has metastasized in a patient;
XX CC (h) assessing the aggressiveness or indolence or prostate cancer in a
XX CC patient;
XX CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX XX
XX XX Sequence 358 BP; 90 A; 85 C; 92 G; 87 T; 0 other:
XX XX
XX XX Query Match 10.7%; Score 41.4; DB 23; Length 358;
XX XX Best Local Similarity 67.6%; Pred. No. 0.0077;
XX XX Matches 71; Conservative 0; Mismatches 33; Indels 1; gaps
XX XX
XX Y 263 CTGAAGTGTACATAAGATGCTTGTCCTTCATACACTCTCGAAGCTGAGATTCGTATTA 342
XX Y 111111 11 1111 11 111111 111111 1 1111 1
XX Db 122 CTGAAGCATCTCCACAGATGATGCTACTGATATTAACCTCCATACAGCTGAGGCTATTA 181
XX Y 343 AATCCTTAGAAGAGAGATGCTCCCTGATGCTATTAAGATATATGCT 367
XX Y 1111111 1 1111 11111111 1111 11 11111
XX Db 182 AATCCTTAGAGAGAGAT-CCCTGAGTCCGTAAGGTATATGCT 225
XX Y
XX XX
XX XX RESULT 12
XX XX ABV22508
XX XX ABV22508 standard; cDNA; 368 BP.
XX XX
XX XX ABV22508:
XX XX

```

[illegible]



```

XX DE Human prostate expression marker cDNA 23903.
XX KM Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker.
XX KM pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN W0200160860-A2.
XX FD 23-AUG-2001.
XX PE 20-FEB-2001; 2001WO-0505171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-185862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-215007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE.
XX DR WPI: 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 4060; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (1) is useful for:
XX CC (a) assessing whether a patient is afflicted with prostate cancer;
XX CC (b) monitoring the progression of prostate cancer in a patient;
XX CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) determining whether prostate cell carcinogenic potential of a compound;
XX CC (g) determining whether prostate cancer has metastasized in a patient;
XX CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
XX CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SO Sequence 388 BP; 94 A; 77 C; 100 G; 113 T; 4 other:
XX
XX Query Match 10.4%; Score 40.4; DB 23; Length 368;
XX Best Local Similarity 66.1%; Pred. No. 0.016;
XX Matches 74; Conservative 0; Mismatches 36; Indels 2; Gaps 1;
XX
XX QY 255 TTGAGTCTCTCTTGAGGAGAAG--CTGAGTCTACTATAAGATCTGTGCTG 31-
XX DB 93 TTGAGATTCTCTCTCAAGGCTCAAGGCTCAAGATCCACAGATATATCTGTG 15-
XX
XX QY 313 ATACTCTAGAGAGTCGACATCTGTATAATCTTAAAGAGGATATCC 364
XX DB 153 ATAACTCCATAGTGTGCGCCCTATACAAATCTCTTCCAGAGAGATCC 204
XX
XX RESULT 14
XX ABV28742
XX ID ABV28742 standard; cDNA; 388 BP.
XX AC ABV28742;
XX AA 10-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 28733.

```

```

XX DE Human prostate expression marker cDNA 13719.
XX KM Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker.
XX KM pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN W0200160860-A2.
XX FD 23-AUG-2001.
XX PE 20-FEB-2001; 2001WO-0505171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-185862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-215007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE.
XX DR WPI: 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 6037; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (1) is useful for:
XX CC (a) assessing whether a patient is afflicted with prostate cancer;
XX CC (b) monitoring the progression of prostate cancer in a patient;
XX CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) determining whether prostate cell carcinogenic potential of a compound;
XX CC (g) determining whether prostate cancer has metastasized in a patient;
XX CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
XX CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SO Sequence 388 BP; 94 A; 77 C; 100 G; 113 T; 4 other:
XX
XX Query Match 10.4%; Score 40.4; DB 23; Length 368;
XX Best Local Similarity 66.1%; Pred. No. 0.016;
XX Matches 74; Conservative 0; Mismatches 36; Indels 2; Gaps 1;
XX
XX QY 255 TTGAGTCTCTCTTGAGGAGAAG--CTGAGTCTACTATAAGATCTGTGCTG 31-
XX DB 93 TTGAGATTCTCTCTCAAGGCTCAAGGCTCAAGATCCACAGATATATCTGTG 15-
XX
XX QY 313 ATACTCTAGAGAGTCGACATCTGTATAATCTTAAAGAGGATATCC 364
XX DB 153 ATAACTCCATAGTGTGCGCCCTATACAAATCTCTTCCAGAGAGATCC 204
XX
XX RESULT 15
XX ABV13728
XX ID ABV13728 standard; cDNA; 382 BP.
XX AC ABV13728;
XX AA 10-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 13719.
XX DE Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

```

pharmacogenomic marker; gene; ss.

homo sapiens.

W0200160860-A2.

23-AUG-2001.

20-FEB-2001: 2001MO-US05171.

17-FEB-2000: 2000US-183319P.

16-MAR-2000: 2000US-189862P.

25-MAY-2000: 2000US-207454P.

09-JUN-2000: 2000US-211314P.

18-JUL-2000: 2000US-219007P.

13-DEC-2000: 2000US-255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI: 2001-662795/76.

Claim 1: Page 2279; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62313) of the

specification or its complement. (I) is useful for:

(a) assessing whether a patient is afflicted with prostate cancer;

(b) monitoring the progression of prostate cancer in a patient;

(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

(e) selecting a composition for inhibiting prostate cancer in a patient;

(f) assessing the prostate cell carcinogenic potential of a compound;

(g) determining whether prostate cancer has metastasized in a patient;

(h) assessing the aggressiveness or indolence of prostate cancer in a patient;

(i) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 392 BP: 96 A; 77 C; 102 G; 114 T; 3 other:

Query Match 10.4%; Score 40.4; DB 23; Length 392;

Best Local Similarity 66.1%; Pred. No. 0.016; Mismatches 36; Indels 2; Gaps 1;

Matches 74; Conservative 0; Mismatches 36; Indels 2; Gaps 1;

255 TTGAGATCATCTTCTTGGAGGAAGAG--CTGAGTGTCTAATAGATCTTGTGCTTC 312

97 TTGAGATCTTCTTCTTGGAGGAAGAG--CTGAGTGTCTAATAGATCTTGTGCTTC 158

313 ATAGCTCTGAGAGCTGAGATCTTGTATTAATCTTGAAGAGAGCATCC 364

157 ATAGCTCTGAGAGCTGAGATCTTGTATTAATCTTGAAGAGAGCATCC 208

Search completed: January 2, 2003, 13:48:57

Job time : 232 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

008 nucleic - nucleic search, using sw model

run on: January 2, 2003, 13:11:27 : Search time 1945 seconds  
(without alignments)  
3224.140 Million cell updates/sec

Title: US-09-762-027-4  
Perfect score: 388  
Sequence: 1 agccgcacatagcaaacat.....atccataatgatalatgcy jbb

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 805774376 residues  
Total number of hits satisfying chosen parameters: 32408132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estopl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	60.6	786	17	A0750247 HS-5576_A
2	138.8	35.8	720	12	BF214297 601848519
3	105.4	27.2	814	12	BF240961 601875551
4	47.6	12.3	709	12	BG430720 602495681
5	45.4	11.7	914	12	BG430555 602504206
6	43	11.1	376	9	AA340632 EST145952

1	502	10	61533471	61533471
2	487	10	61533471	61533471
3	487	10	61533471	61533471
4	487	10	61533471	61533471
5	487	10	61533471	61533471
6	487	10	61533471	61533471
7	487	10	61533471	61533471
8	487	10	61533471	61533471
9	487	10	61533471	61533471
10	487	10	61533471	61533471
11	487	10	61533471	61533471
12	487	10	61533471	61533471
13	487	10	61533471	61533471
14	487	10	61533471	61533471
15	487	10	61533471	61533471
16	487	10	61533471	61533471
17	487	10	61533471	61533471
18	487	10	61533471	61533471
19	487	10	61533471	61533471
20	487	10	61533471	61533471
21	487	10	61533471	61533471
22	487	10	61533471	61533471
23	487	10	61533471	61533471
24	487	10	61533471	61533471
25	487	10	61533471	61533471
26	487	10	61533471	61533471
27	487	10	61533471	61533471

## ALIGNMENTS

RESULTS  
LOCUS: 576-A2-E06-SPR-KR1-11 Human Male BAC library; Homo sapiens  
DEFINITION: A1723653  
VERSION: 1  
KEYWORDS: BAC, 576-A2-E06-SPR-KR1-11 Human Male BAC library; Homo sapiens  
ADDITIONAL INFORMATION: 576-A2-E06-SPR-KR1-11 Human Male BAC library; Homo sapiens  
SOURCE: GenBank  
ORGANISM: Homo sapiens  
REFERENCE: 1. Smith, K., Swartzell, S., Holzman, T., Waller, E., Butler, J., Young, J., Zhai, S., Adams, M.D. and Miller, R. (1999) A sequence approach to mapping and defining the human genome  
2. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
3. Human Genome Project  
4. Human Genome Project  
5. Human Genome Project  
6. Human Genome Project  
7. Human Genome Project  
8. Human Genome Project  
9. Human Genome Project  
10. Human Genome Project  
11. Human Genome Project  
12. Human Genome Project  
13. Human Genome Project  
14. Human Genome Project  
15. Human Genome Project  
16. Human Genome Project  
17. Human Genome Project  
18. Human Genome Project  
19. Human Genome Project  
20. Human Genome Project  
21. Human Genome Project  
22. Human Genome Project  
23. Human Genome Project  
24. Human Genome Project  
25. Human Genome Project  
26. Human Genome Project  
27. Human Genome Project

COMMENT: 1. Smith, K., Swartzell, S., Holzman, T., Waller, E., Butler, J., Young, J., Zhai, S., Adams, M.D. and Miller, R. (1999) A sequence approach to mapping and defining the human genome  
2. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
3. Human Genome Project  
4. Human Genome Project  
5. Human Genome Project  
6. Human Genome Project  
7. Human Genome Project  
8. Human Genome Project  
9. Human Genome Project  
10. Human Genome Project  
11. Human Genome Project  
12. Human Genome Project  
13. Human Genome Project  
14. Human Genome Project  
15. Human Genome Project  
16. Human Genome Project  
17. Human Genome Project  
18. Human Genome Project  
19. Human Genome Project  
20. Human Genome Project  
21. Human Genome Project  
22. Human Genome Project  
23. Human Genome Project  
24. Human Genome Project  
25. Human Genome Project  
26. Human Genome Project  
27. Human Genome Project

1. Smith, K., Swartzell, S., Holzman, T., Waller, E., Butler, J., Young, J., Zhai, S., Adams, M.D. and Miller, R. (1999) A sequence approach to mapping and defining the human genome  
2. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
3. Human Genome Project  
4. Human Genome Project  
5. Human Genome Project  
6. Human Genome Project  
7. Human Genome Project  
8. Human Genome Project  
9. Human Genome Project  
10. Human Genome Project  
11. Human Genome Project  
12. Human Genome Project  
13. Human Genome Project  
14. Human Genome Project  
15. Human Genome Project  
16. Human Genome Project  
17. Human Genome Project  
18. Human Genome Project  
19. Human Genome Project  
20. Human Genome Project  
21. Human Genome Project  
22. Human Genome Project  
23. Human Genome Project  
24. Human Genome Project  
25. Human Genome Project  
26. Human Genome Project  
27. Human Genome Project

(pfeier@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering/bacpac.htm>) or from Research Genetics (info@resgen.com). BAC end Web Server: <http://www.husc.washington.edu>  
 Plate: 1152 row: 1 column: 16  
 Seq primer: SP6  
 Class: BAC ends  
 High quality sequence stop: 786.

FEATURES

Source  
 Location/Qualifiers  
 1..786  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="Plate 1152 Col 16 Row 1"  
 /clone\_lib="KPI-11 Human Male BAC Library"  
 /sex="male"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"  
 134 c 165 g 222 t 2 others

Query Match  
 Best Local Similarity 94.08; Score 235; Db 17; Length 786;  
 Matches 266; Conservative 0; Mismatches 15; Indels 2; Gaps -;

1 ANCTGCTCATATGACGACATATTCAGTCCTCTCTGCTCTCAAGGCTGATTTCT 5'  
 |||||  
 405 AGCTGCGCATATGACGACATATTCAGTCCTCTCTGCTCTCAAGGCTGATTTCT 345  
 61 TGCTATGACGACAGCTGCTGTTATCTACGAGAACCCATACCTCTGAGCTTTAT 140  
 |||||  
 345 TGCTATGACGACAGCTGCTGTTATCTACGAGAACCCATACCTCTGAGCTTTAT 285  
 121 GCTTACAGGACAGCTGAGTGAATTTATGCTTTCAGTCCGTACAGTTATTA 180  
 |||||  
 285 GCTTACAGGACAGCTGAGTGAATTTATGCTTTCAGTCCGTACAGTTATTA 226  
 181 GCTTACAGGACCTTTCATTTGCCATCTGAGAACTGAGGAGGAGGAGTATAC 240  
 |||||  
 225 GCTTACAGGACCTTTCATTTGCCATCTGAGAACTGAGGAGGAGGAGTATAC 175  
 241 GTTACAGGACCTTTCATCTGCTTTCAGGAGGAGGAGTATAC 263  
 |||||  
 167 GTTACAGGACCTTTCATCTGCTTTCAGGAGGAGGAGTATAC 125

RESULT 2  
 BE214297  
 LOCUS 60184851961 NIH\_MGC\_55 Homo sapiens cDNA clone IMAGE:409238 5'  
 DEFINITION mRNA sequence.  
 ACCESSION BE214297  
 VERSION BE214297.1 GI:11107883  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;  
 1 (bases 1 to 720)  
 NIH-MGC <http://mgc.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATOC  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLMF at:  
<http://image.jhl.gov>

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 Source  
 Location/Qualifiers  
 1..500  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="Plate 1152 Col 16 Row 1"  
 /clone\_lib="KPI-11 Human Male BAC Library"  
 /sex="male"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"  
 134 c 165 g 222 t 2 others

Query Match  
 Best Local Similarity 94.08; Score 235; Db 17; Length 786;  
 Matches 266; Conservative 0; Mismatches 15; Indels 2; Gaps -;

1 ANCTGCTCATATGACGACATATTCAGTCCTCTCTGCTCTCAAGGCTGATTTCT 5'  
 |||||  
 405 AGCTGCGCATATGACGACATATTCAGTCCTCTCTGCTCTCAAGGCTGATTTCT 345  
 61 TGCTATGACGACAGCTGCTGTTATCTACGAGAACCCATACCTCTGAGCTTTAT 140  
 |||||  
 345 TGCTATGACGACAGCTGCTGTTATCTACGAGAACCCATACCTCTGAGCTTTAT 285  
 121 GCTTACAGGACAGCTGAGTGAATTTATGCTTTCAGTCCGTACAGTTATTA 180  
 |||||  
 285 GCTTACAGGACAGCTGAGTGAATTTATGCTTTCAGTCCGTACAGTTATTA 226  
 181 GCTTACAGGACCTTTCATTTGCCATCTGAGAACTGAGGAGGAGGAGTATAC 240  
 |||||  
 225 GCTTACAGGACCTTTCATTTGCCATCTGAGAACTGAGGAGGAGGAGTATAC 175  
 241 GTTACAGGACCTTTCATCTGCTTTCAGGAGGAGGAGTATAC 263  
 |||||  
 167 GTTACAGGACCTTTCATCTGCTTTCAGGAGGAGGAGTATAC 125

RESULT 3  
 BE214297  
 LOCUS 60184851961 NIH\_MGC\_55 Homo sapiens cDNA clone IMAGE:409238 5'  
 DEFINITION mRNA sequence.  
 ACCESSION BE214297  
 VERSION BE214297.1 GI:11107883  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;  
 1 (bases 1 to 720)  
 NIH-MGC <http://mgc.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATOC  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLMF at:  
<http://image.jhl.gov>

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 Source  
 Location/Qualifiers  
 1..500  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="Plate 1152 Col 16 Row 1"  
 /clone\_lib="KPI-11 Human Male BAC Library"  
 /sex="male"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"  
 134 c 165 g 222 t 2 others







Best Local Similarity 68.6% Pred. No. 0.071:  
Matches 72: Conservative 0: Mismatches 32: Indels 1: Gaps 1:

QY 263 CTGAAGTCTACATAGATGCTTGCTTTCATACCTCAGAAAGTCCAGATTCCTGA 342  
111111 11 111 11 111111 1 11111 1 1111  
DB 378 CTGAAGCATCCAGCAAGATGCTTACGATGATACCTCCCAACAGATGCTGCGCATATCA 437  
343 AATCCTTAGAAAAGCAGATCCCTCAATCATTAANGATATATG 367  
111111 1 11111 11111 111 1 11 11 11 11  
DB 438 AATCCTTGCCAGAGCAT-CCCTGAGTCCCTGAGAGGATATGCT 481

RESULT 11  
LOCUS B1764628 688 bp mRNA linear EST 25-SEP-01  
DEFINITION 603050942P1 NIH-MGC\_116 Homo sapiens cDNA clone IMAGE:15152675 5'  
mRNA sequence.  
B1764628  
B1764628.1 GI:15756206  
EST.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;  
1 (bases 1 to 688)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
RNA Sequencing by: Incyte Genomics, Inc.  
cDNA distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: L1AM11477 row: d column: 12  
High quality sequence stop: 686.  
Location/Qualifiers

FEATURES  
Source

1..688  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5190875"  
/clone\_lib="NIH-MGC\_116"  
/lab\_host="DH10B"  
/note="Organ: pooled colon, kidney, stomach; Vector:  
pCMV-SPORT6; Site:1: Ncol; Site:2: EcoRV (destroyed); RNA  
source anonymous pool of 3 clones; age 26 yo male, 49 yo  
female, 71 yo male colon; 46 yo female kidney, and pool of  
stomachs, 62 yo male and 70 yo female. Library is  
oligo-dt primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.4 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by  
Gruber (Invitrogen). Research Genetics tracking code  
023. Note: this is a NIH-MGC Library."

BASE COUNT 191 a 148 c 157 g 192 t  
ORIGIN

Query Match 11.1% Score 43: DB 13: Length 688:  
Best Local Similarity 68.6% Pred. No. 0.071:  
Matches 72: Conservative 0: Mismatches 32: Indels 1: Gaps 1:

QY 263 CTGAAGTCTACATAGATGCTTGCTTTCATACCTCAGAAAGTCCAGATTCCTGA 342  
111111 11 111 11 111111 1 11111 1 1111  
DB 378 CTGAAGCATCCAGCAAGATGCTTACGATGATACCTCCCAACAGATGCTGCGCATATCA 437  
343 AATCCTTAGAAAAGCAGATCCCTCAATCATTAANGATATATG 367  
111111 1 11111 11111 111 1 11 11 11 11  
DB 438 AATCCTTGCCAGAGCAT-CCCTGAGTCCCTGAGAGGATATGCT 481

Best Local Similarity 70.2% Pred. No. 0.071:  
Matches 72: Conservative 0: Mismatches 32: Indels 1: Gaps 1:

QY 263 CTGAAGTCTACATAGATGCTTGCTTTCATACCTCAGAAAGTCCAGATTCCTGA 342  
111111 11 111 11 111111 1 11111 1 1111  
DB 378 CTGAAGCATCCAGCAAGATGCTTACGATGATACCTCCCAACAGATGCTGCGCATATCA 437  
343 AATCCTTAGAAAAGCAGATCCCTCAATCATTAANGATATATG 367  
111111 1 11111 11111 111 1 11 11 11 11  
DB 438 AATCCTTGCCAGAGCAT-CCCTGAGTCCCTGAGAGGATATGCT 481

RESULT 11  
LOCUS B1764628 702 bp mRNA linear EST 11-MAR-2001  
DEFINITION 603050942P1 NIH-MGC\_75 Homo sapiens cDNA clone IMAGE:1608746 5'  
mRNA sequence.  
B1764628  
B1764628.1 GI:15756206  
EST.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;  
1 (bases 1 to 702)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: CLONTECH Laboratories, Inc.  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
RNA Sequencing by: Incyte Genomics, Inc.  
cDNA distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: L1AM1351 row: e column: 03  
High quality sequence stop: 693.  
Location/Qualifiers

1..702  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1608746"  
/clone\_lib="NIH-MGC\_75"  
/lab\_host="DH10B (TI phage-resistant)"  
/note="Organ: kidney; Vector: pDR-LIB (Clontech); Site:1:  
SfiI (ggccatcgccg); Site:2: SfiI (ggccatcgcc); 5' and  
3' adaptors were used in cloning as follows: 5' adaptor  
sequence: 5'-CACGGCCATATGCGC-3' and 3' adaptor sequence:  
5'-ATTCAGAGCGCGCGCCGAGATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.65  
kb (range 0.5-1.0 kb). 15/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH-MGC Library."

BASE COUNT 145 a 145 c 196 g 173 t  
ORIGIN

Query Match 11.1% Score 43: DB 12: Length 702:  
Best Local Similarity 68.6% Pred. No. 0.071:  
Matches 72: Conservative 0: Mismatches 32: Indels 1: Gaps 1:

QY 263 CTGAAGTCTACATAGATGCTTGCTTTCATACCTCAGAAAGTCCAGATTCCTGA 342  
111111 11 111 11 111111 1 11111 1 1111  
DB 378 CTGAAGCATCCAGCAAGATGCTTACGATGATACCTCCCAACAGATGCTGCGCATATCA 437  
343 AATCCTTAGAAAAGCAGATCCCTCAATCATTAANGATATATG 367  
111111 1 11111 11111 111 1 11 11 11 11  
DB 438 AATCCTTGCCAGAGCAT-CCCTGAGTCCCTGAGAGGATATGCT 481



AUTHORS NIH-MGC <http://imgc.ncl.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMC)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1CM1330 row: f column: 09  
 High quality sequence stop: 695.  
 Location/Qualifiers

FEATURES  
 source  
 1..703  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4592264"  
 /lab\_host="NIH-MGC-75"  
 /lab\_host="DH10B (11 phage-resistant)"  
 /note="Organ: Kidney; Vector: pMR-11b (Clontech); Site: 1; Site 1 (ggccgcctggcc); Site 2: Site 1 (ggccgcctggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATTAATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGCCGCCGACATG-dt(30)BN-3' (where B = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

BASE COUNT  
 203 a 151 c 168 g 181 t

Query Match 11.18; Score 43; DB 12; Length 703;  
 Best Local Similarity 68.6%; Pred. No. 0.074;  
 Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

263 CTGAGTGTCTCAAGTGGTGTCTGCTCACTCAAGTCAAGTCTGATAT 342  
 241 CTGAGGATCCACAGATGATCTCTGATGATATCTCCATAGCTGTGCCCTATACA 300

343 AATCCTTAGAAGAGCATCCCTGGAATCCATAAAGTATATATGCG 367  
 301 AATCCTTGCCAGAGGATCCCTGAGTCCCTGAGAGGTATATGCC 344

RESULT 14  
 B3400739 734 bp mRNA linear EST 12-MAR-2001  
 LOCUS 602461127F1 NIH-MGC\_75 Homo sapiens cDNA clone IMAGE:4592264 5'  
 DEFINITION mRNA sequence.  
 ACCESSION B3400739  
 VERSION B3400739  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NIH-MGC <http://imgc.ncl.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1CM1330 row: e column: 08

REFERENCE  
 AUTHORS NIH-MGC <http://imgc.ncl.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1CM1330 row: e column: 08

High quality sequence stop: 726.  
 Location/Qualifiers  
 1..744  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4592264"  
 /lab\_host="NIH-MGC-75"  
 /lab\_host="DH10B (11 phage-resistant)"  
 /note="Organ: Kidney; Vector: pMR-11b (Clontech); Site: 1; Site 1 (ggccgcctggcc); Site 2: Site 1 (ggccgcctggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATTAATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGCCGCCGACATG-dt(30)BN-3' (where B = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

BASE COUNT  
 210 a 153 c 179 g 185 t

Query Match 11.18; Score 43; DB 12; Length 744;  
 Best Local Similarity 68.6%; Pred. No. 0.076;  
 Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

263 CTGAGTGTCTCAAGTGGTGTCTGCTCACTCAAGTCAAGTCTGATAT 342  
 241 CTGAGGATCCACAGATGATCTCTGATGATATCTCCATAGCTGTGCCCTATACA 300

343 AATCCTTAGAAGAGCATCCCTGGAATCCATAAAGTATATATGCG 367  
 301 AATCCTTGCCAGAGGATCCCTGAGTCCCTGAGAGGTATATGCC 344

RESULT 15  
 B3428541 744 bp mRNA linear EST 14-MAR-2001  
 LOCUS 602501062F1 NIH-MGC\_75 Homo sapiens cDNA clone IMAGE:4514667 5'  
 DEFINITION mRNA sequence.  
 ACCESSION B3428541  
 VERSION B3428541  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NIH-MGC <http://imgc.ncl.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1CM1330 row: f column: 16  
 High quality sequence stop: 706.  
 Location/Qualifiers

FEATURES  
 source  
 1..744  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4514667"  
 /lab\_host="NIH-MGC-75"  
 /lab\_host="DH10B (11 phage-resistant)"  
 /note="Organ: Kidney; Vector: pMR-11b (Clontech); Site: 1; Site 1 (ggccgcctggcc); Site 2: Site 1 (ggccgcctggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATTAATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGCCGCCGACATG-dt(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

BASE COUNT 203 a 159 c 169 g 212 t 1 others

Query Match 11.18; Score 43; DB 12; Length 744;

Best Local Similarity 68.68; Pred. NO. 0.076; Mismatches 32; Indels 1; Gaps 1;

Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 283 CTGAGTGTCTACATPAAGATCTGTCCTTCATPACTCTCAGAGGTCAGATTCCTGATA 342

Db 382 CTGAGCATCTCCACAGATGATCTCTACTGAATTAATCCCATAGCTGCTGCGCCCTATACA 441

QY 343 AATCCTTAGAAAGAGCATCCCTGATCCATCAATAGATATATGCG 387

Db 442 AATCCTTGCCAGAGGAGCAT-CCCTGAGTCCCTGAGAGGTGATAGCC 485

Search completed: January 2, 2003, 14:21:43

Job time : 1955 secs



Fri Jan 3 14:27:02 2003

us-09-762-027-4.rni

Page 2

```

APPLICANT: Yue, Henry
APPLICANT: Guejler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 1249 base pairs
TYPE: nucleic acid
STRANDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SPLN0702
CLONE: 208836
US-08-933-750C-62

Query Match 8.0%: Score 31; DB 2; Length 1249;
Best Local Similarity 57.3%: PctId No. 2;
Matches 55; Conservative 0; Mismatches 41; Indels 0; Gaps

CY 288 GTCCACATAGAGCTGTGCGCTTCATCAATCTCAGAAAGCGTAGAGATTGTAATATATTC 317
DB 1122 GAGCAGCAGGAGGAGGAGGAGGCTGCGACTGCGCAGAGAGCCCGATGAGAGTGCATTC 1153
CY 348 TTAGAAAAGACATCCCTCGATCCATCAATGAAGTATA 383
DB 1182 CCATCTAGCTAGCTTCCCATTAATGCTTAATTTAAA 1217

RESULT 3
US-09-234-613-62
Sequence 62, Application US/09234613
Patent No. 6132973
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guejler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.

```

[illegible]





```

ADDRESS: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple System 7.5.3
SOFTWARE: Microsoft Word, Version 6.0.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,258
FILING DATE: 26-SEPT-1996
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: US 08/575,715
FILING DATE: 20-DEC-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,324
FILING DATE: 29-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2620-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 758 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: Cowpox p35
FEATURE:
NAME/KEY: CDS
LOCATION: 18..758
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 66..755
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 18..65
US-08-720-258-1

Query Match 7.6%; Score 29.4; hb 4; Length 758;
Best Local Similarity 58.6%; Pred. No. 5.1;
Matches 51; Conservative 0; Mismatches 36; Inlets 0; Ips

QY 216 CTGGCAGCAGCAGTGTCCTACGAGCTGTCACGAAGAACTTCAGATCATCTCTTCAAGC 273
||||| ||| | | ||||| |||| | | || | || | || | || | || | || | || |
DB 48 CTCGCGCAGAGTTCATATGCCATACGACGCTTCACGACATCATCTCTCATCTCTCAT 177
||||| ||| | | ||||| |||| | | || | || | || | || | || | || | || |

QY 276 AAGAACGCTGAAGTGTCTACATATG 302
| |||| ||| |||| | |||
DB 108 ACGGAAGAAGAAACACACATCATATG 134
| |||| ||| |||| | |||

RESULT 11
US-08-381-691-17
Sequence 17, Application US/08381691
Patent No. 5852224
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Alpha-lac Albumin Gene Constructs
NUMBER OF SEQUENCES: 17
COMPUTER READABLE FORM:

```

```

MEDICAL APT: 11/17/1988
REPORT: IBM PC COMPATIBLE
OPERATING SYSTEM: DOS/MS-DOS
SUBJECT: PATENT IN REPLY TO: 11/17/1988
CURRENT APPLICATION DATA: US/06/01/1691
APPLICATION NUMBER: US/06/01/1691
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2119 base pairs
TYPE: nucleic acid
STRATEGY: double
POLYMER: linear
Molecule TYPE: cDNA
HYPERMUTATION: NO
ANTI-SENSE: NO
US/06/01/1691-17

Query Match: 7 58: Score 29.4: DB 2: Length 2119:
Best Local Similarity: 51.18: Pctd. No. 7.9:
Matches 09: Conservative 0: Mismatches 66: Indels 0: Gaps 0:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 9
```









Fri Jan 3 14:27:03 2003

us-09-762-027-4.rnpb

Page 2

```

APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Foad.
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: NO. US20020082206A1el Polynucleotides from Altreigen "S...
FILE OF INVENTION: Thereaby
FILE REFERENCE: 21402-013 (Cura-313)
CURRENT APPLICATION NUMBER: US/05/867,550
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: USSN 60/208,427
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: PASTSEQ for Windows Version 4.0
SEQ ID NO: 925
LENGTH: 755
TYPE: DNA
ORGANISM: Homo sapiens
5'-05-867-350-929

```

	Query Match	Similarity	68.6%	Score 47	DB 10;	Length 797		
	Best Local			Pred. No.	0.0092			
	Matches	Conservative	0	Mismatches	33	Indels	1	Refs
QY	263	CIGAGTCTCATATGAGTAGTGCTGTGGCTTCATAATCCGAAGCTGTAACATTTAAAC	312					
LQ	95	CTGAAGCATTCCCACAGATGATGCCTACTCGAATAAATGCCATTAACCTCGAGGCACTAA	154					
LY	343	AATCTCTTAGAAAGAAGCATCCCCPCAATCCATAAANNITAAIANNLNC	387					
LQ	155	AATCTCTTGCCGACAGACAT-CCCTAGATGCCCTGAAGGCTGIATGACC	198					

```

1 RESULT 3
2 US-09-833-381-1948
3 : Sequence 1948, Application US/09833381
4 : Patent No. US20020132090A1
5 :
6 : GENERAL INFORMATION:
7 :
8 : APPLICANT: Robison, Keith E.
9 : TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
10 : FILE REFERENCE: 5800-119
11 : CURRENT APPLICATION NUMBER: US/09/833,381
12 : CURRENT FILING DATE: 2001-04-11
13 : PRIORITY APPLICATION NUMBER: 09/516,448
14 : PRIORITY FILING DATE: 2000-02-29
15 :
16 : NUMBER OF SEQ. ID NOS: 2050
17 :
18 : SOFTWARE: FASTSEQ for Windows Version 3.0
19 : SEQ. ID NO 1948
20 :
21 : LENGTH: 563
22 :
23 : TYPE: DNA
24 :
25 : ORGANISM: Homo sapiens
26 :
27 : FEATURE:
28 : NAME/KEY: misc_feature
29 : LOCATION: (1)..(963)
30 :
31 : OTHER INFORMATION: n = A,T,C or G
32 :
33 : US-09-833-381-1948

```

Query Match	11.1%	Score 43	DB 101	Length 363
Best Local similarity	68.6%	Pred. No. 0.001		
Matches	72: Conservative	0: Mismatches	32: Indels	1: Gaps
283	CTGAAGCTCTCATATAGATGCTGTGTGCTTCCTCAATATCTCAGAACGCTGGAGATTCCTGATA	342		
337	CTGAAGCATCCACAGATGATGATCTACTGATGATATCTCCATATAGGCTGGCTGATATAT	396		
343	AATCTCTTAGAAAGAGCATCCCTCGATATCCATAAANGATATATNG	367		
397	AATCTCTTAGCCAGAGCAT-CCCTTAGTCTCCCTGAAGCTGTATGGC	440		

RESULT 4  
US-05-925-300-125  
: Sequence 125, Application US/09925300

SELECT \* FROM   
 APPLICANT WHERE   
 APPLICANT STATE = "CA"   
 TITLE = "INVENTOR OF PATENT RIGHTS, PROTEINS AND ANTIBODIES"   
 FILE NUMBER = "1111"   
 CURRENT APPLICANT'S NAME = "S/09/225,300"   
 CURRENT FILING DATE = "7/10/2010"   
 PRIOR APPLICANT'S NUMBER = "1/0500/45988"   
 PRIOR FILING DATE = "2/09/2007"   
 PRIOR APPLICANT'S NUMBER = "7/24,270"   
 PRIOR FILING DATE = "1/27/2007"   
 NUMBER OF SHEETS = "1"   
 STEWARD FULTON, YOUNG &   
 SELLER'S 125   
 LENDER: 1189   
 IPER: BNA   
 ROASTSM: dm: sh:ios   
 FEATDF:   
 NAME/REV: m: f: r: o:   
 LOCATION: (1141)   
 OTHER INFORMATION: no reference to   
 NAME/REV: m: f: r: o:   
 LOCATION: (1126)   
 OTHER INFORMATION: no reference to   
 NAME/REV: m: f: r: o:   
 LOCATION: (1144)   
 OTHER INFORMATION: no reference to   
 S/09/225-300-125

[illegible]

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

Query Match:	9.14	Score 37.6;	E6 10;	Length 195;
Best Local Similarity:	64.44	Prod. NO. 0.023;		
Matches	75;	Conservative	0;	Mismatches 30;
			Indels	0;
			Gaps	0;

QY	280	AAGCTGAGGTCTGCACATPAGATGTCGTGGCTCAACTCTCCAGAACCGTAGATT	167	33
Dδ	56	AAGCTTACCTGTCTCNCATCGATGATGTCTCTCCCTCCCAAGTGCCCATGTCGTAAGTCTGG	157	
QY	340	ATAAATCCTTAGAAAAAGAGCATGCC	364	
Dδ	158	AGAATACTCTTAGAGAAAGACCCTTCC	182	

```

: RESULT 6
: US-05-960-352-7792
: Sequence 7792, Application US/05960352
: Patent No. US20020137135A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Nengping
: APPLICANT: Byatt, John C.
: APPLICANT: Mathialagan, Nagappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
: FILE REFERENCE: 16511.006/37-21(10298)C
: CURRENT APPLICATION NUMBER: US/09/960,352
: CURRENT FILING DATE: 2001-05-24
: NUMBER OF SEQ ID NOS: 15112
: SEQ ID NO 7792
: LENGTH: 313
: TYPE: DNA
: ORGANISM: bos taurus
: OTHER INFORMATION: Clone ID: 33-LIB34-076-Q1-E1-A2
: US-05-960-352-7792

```

	Query Match	9.5%;	Score 37;	DB 10;	Length 319;	--
	Best Local Similarity	64.7%;	Pred. NO.	0.044		
Matches	55; Conservative	0;	Mismatches	30;	Indels	0; Gaps
CY	280 AACCTAAGTCACACTAAATGATCGTGTCTGCATCTCAATCCTCTCATGAAGAAGTGTAATTCTGT	359				
TDb	58 AAGCTTACTGTCGTCTTCATCATGATTTCTCTCTCTCAAGTCCCGATATCTCTCTGGATCTCTGG	157				
CY	340 ATAAATCCTTAGAAAAGACATGCC	364				
EL	156 AGAATACTCTTAGAAGAACGCTTCC	182				

```

1      RESULT 7
2      US-09-960-352-14120
3      : Sequence 14120, Application US/09560352
4      : Patent No. US20020137139A1
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: Warren, Wesley C.
9      : APPLICANT: Tao, Mengping
10     : APPLICANT: Byatt, John C.
11     : APPLICANT: Mathialagan, Nagappan
12     :
13     : TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
14     :
15     : TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
16     :
17     : FILE REFERENCE: 16511.006/37-21(10298)C
18     :
19     : CURRENT APPLICATION NUMBER: US/09/960.352
20     :
21     : CURRENT FILING DATE: 2001-09-24
22     :
23     : NUMBER OF SEQ ID NOS: 15112
24     :
25     : SEQ ID NO 14120
26     :
27     : LENGTH: 303
28     :
29     : TYPE: DNA
30     :
31     : ORGANISM: Bos taurus
32     :
33     : OTHER INFORMATION: Clone ID: 60-LIB34-U27-Q1-E1-G8
34     :
35     : US-09-960-352-14120

```

[illegible]

27	340	6.166666667	6.166666667	324
28	1	1111	1111	325
29	212	4.166666667	4.166666667	326

```

1      RESULT 6
2      US-09-960-11-780
3      :
4      : Sequence ID: Application US/09060352
5      : Patent No. 54001917139A1
6      :
7      : GENERAL INFORMATION:
8      :
9      : APPLICANT: Warren, Wesley C.
10     :
11     : APPLICANT: Bao, Wensheng
12     :
13     : APPLICANT: Byatt, John C.
14     :
15     : APPLICANT: Nathalaagan, Nagappan
16     :
17     : TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
18     :
19     : FILE REFERENCE: 16911.006/37-21(102587C)
20     :
21     : CURRENT APPLICANT NUMBER: US/09/960.352
22     :
23     : CURRENT FILING DATE: 2001-03-24
24     :
25     : NUMBER OF SEQ ID NOS: 15112
26     :
27     : SEQ ID NO 1:
28     :
29     : LENGTH: 17
30     :
31     : TYPE: LNA
32     :
33     : ORGANISM: POS Taurus
34     :
35     : OTHER INFORMATION: Clone ID: 29-11B34-011-Q1-E1-H1
36     :
37     : US-09-960-11-780

```

Query Match	Y	E	Score	37	DB	10	Length	39
Best Local Similarity	64.74		Primal NO	0.049				
Matches	51	Conservative			Mismatches	30	Indels	0
QY	286	AACTTAACTGATACATACAGTCTGCTGCTCATACCTGAGAGCTGCAGATTCTGT	339					
DB	183	AGCTTAAGTGTCTTGCATCATGCTTCTCTGCAAGGTCGCCAGATGCTGCAGATGCTGG	242					
QY	340	ATATATGCTTAAAGAAAGATATCCG	364					
DB	243	AATATGCTTAAAGAAAGATATCCG	267					

```

1  RESULT 9
2  US-05-960-352-770
3  ? Sequence ID: Application US/05960352
4  ? Patent No.: 50026157135A1
5  ? GENERAL INFORMATION:
6  ? APPLICANT: Carter, Wesley C.
7  ? APPLICANT: Iao, Mengbing
8  ? APPLICANT: Byatt, John C.
9  ? APPLICANT: Mathilagan, Nalini
10 ? TITLE OF INVENTION: NOCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
11 ? TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
12 ? FILE REFERENCE: 1651.006/3721(102980C
13 ? CURRENT PUBLICATION NUMBER: US/05/960.352
14 ? CURRENT FILING DATE: 2001-03-24
15 ? NUMBER OF SEQ ID NOS: 13112
16 ? SEQ ID NO: 1
17 ? LENGTH: 111
18 ? TYPE: DNA
19 ? ORGANISM: Bos taurus
20 ? FEATURES:
21 ? NAME/KEY: 3577
22 ? LOCATION: 3577
23 ? OTHER INFORMATION: unsure if all in locations
24 ? OTHER INFORMATION: Clone ID: 29-11834-001-01-E1-H1
25 ? US-05-960-352-770

```

```

Query Match: 9.54; Score 37; DB 10; Length 392;
Best Local Similarity 64.7%; Pred. No. 0.049; 30; Indels 0; Caps 0
Matches 57; Conservative 0; Mismatches 30;
260 AAGTACAGTACATTAATGCTGTCCTGTCACAACTGCGAAGAGCTACAGTCTGT 339
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



QY 340 ATAAATCCTTAGAAGACCATCC 364  
| | | | | | | | | | | | | | | | | |  
Db 223 AGAATCCTTGATGAGAGCCTTCC 247

## RESULT 14

US-09-770-445-209  
Sequence 209, Application US/09770445  
Patent No. US200202281A1

## GENERAL INFORMATION:

APPLICANT: Gorlach, Jorn  
APPLICANT: An, Yong-Qiang  
APPLICANT: Hamilton, Carol M.  
APPLICANT: Price, Jennifer L.  
APPLICANT: Raines, Tracy M.  
APPLICANT: Yu, Yang  
APPLICANT: Rameaka, Joshua G.  
APPLICANT: Page, Amy  
APPLICANT: Mathew, Abraham V.  
APPLICANT: Ledford, Brooke L.  
APPLICANT: Moessner, Jeffrey P.  
APPLICANT: Haas, William David  
APPLICANT: Garcia, Carlos A.  
APPLICANT: Kricke, Ted  
APPLICANT: Slader, Ted  
APPLICANT: Davis, Keith R.  
APPLICANT: Allen, Keith  
APPLICANT: Hoffman, Neil  
APPLICANT: Hurban, Patrick  
TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
FILE REFERENCE: 2023US (PARA-012PRV)  
CURRENT APPLICATION NUMBER: US/09/770,445  
CURRENT FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: US 60/178,472  
PRIOR FILING DATE: 2000-01-27  
NUMBER OF SEQ ID NOS: 999  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 209  
LENGTH: 1022  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana

## FEATURE:

NAME/KEY: misc\_feature  
LOCATION: (1)...(1022)

OTHER INFORMATION: n = A,T,C or G  
US-09-770-445-209

Query Match 8.5%; Score 32.8; DB 10; Length 1022;  
Best Local Similarity 51.4%; Pred. No. 1.6;

Matches 76; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 213 AAATGACAGCAGCAGCTGCTACAGCTCTCAAGAAGACTTCATCTTCTTGA 272  
| | | | | | | | | | | | | | | | | |  
Db 12 AAAGTGTGGAAGATATTATTACACCAAAAAGACATGATTCAGAGTTGG 71

QY 273 GGAAGAAGCTGAAGTGTCTACATAGATGCTGTCTTATAACTCTGAGAGCTTCAG 332  
| | | | | | | | | | | | | | | | | |  
Db 72 AAGACAGAGCTGAAGTGTCTTAAATTAAGATTGTCGAGAGTTTCCTACCAAAACT 131

QY 333 ATTCTGATTAATCTCTAGAAAGAGCA 360  
| | | | | | | | | | | | | | | | | |  
Db 132 TTATTATCAAAATGAAAAACAAAGAA 159

## RESULT 15

US-09-796-692-5227/C

Sequence 5227, Application US/09796692

Publication No. US20020198362A1

## GENERAL INFORMATION:

APPLICANT: Gaigier, Alexander  
APPLICANT: Aigater, Paul A.  
APPLICANT: Mannion, Jane

TITLE OF INVENTION: NEW STRAINS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER

FILE REFERENCE: 2007-06-07  
CURRENT APPLICATION NUMBER: US/09/796,692

PRIOR APPLICATION NUMBER: 60/185,126

PRIOR FILING DATE: 2000-07-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-07-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-07-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: 60/200,599

PRIOR FILING DATE: 2000-07-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-07-04

PRIOR APPLICATION NUMBER: 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: 60/216,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/229,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: 60/223,376

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9797

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 527

LENGTH: 533

TYPE: DNA

ORGANISM: H. sapiens

US-09-796-692-5227

Query Match 5.3%; Score 32.2; DB 9; Length 533;  
Best Local Similarity 52.6%; Pred. No. 1.8;

Matches 70; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 32 TCTCTTCTGTAAGAGTGTGCTGCTATGACAGACAGCTGCTTATCTA 51  
| | | | | | | | | | | | | | | | | |  
Db 339 TCTCTTCTGTAAGAGTGTGCTGCTATGACAGACAGCTGCTTATCTA 280

QY 92 GAGAGTGTGTAAGAGTGTGCTGCTATGACAGACAGCTGCTTATCTA 151  
| | | | | | | | | | | | | | | | | |  
Db 279 AAGAGTGTGTAAGAGTGTGCTGCTATGACAGACAGCTGCTTATCTA 220

QY 152 TTATTATCAAAATGAAAAACAAAGAA 159  
| | | | | | | | | | | | | | | | | |  
Db 219 TTATTATCAAAATGAAAAACAAAGAA 159

Search completed: January 4, 2003, 15:07:16  
Job time: 53 s

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

UM nucleic - protein search, using frame\_plus.n2p model

Run on: January 2, 2003, 15:07:19 ; search time 42 seconds

(without alignments)  
3390.923 Million cell updates/sec

Title: US-09-762-027-4

Perfect score: 657  
Sequence: 1 agctgcataatcagcaacat.....atccataaattatattatg aac

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Dgapop 6.0	Dgapext 7.0
Delop 6.0	Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 56446

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+.n2p.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US09762027/rnart.02012003\_085613\_26c2/app.query.fasta...563  
-DB=PIR\_73 -QEM=fastan -SUFFIX=20.rpr -MINMATCH=0.1 -LOCEXT=0 -LOCEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human.ucali -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFM=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09762027.fcen.1-1-16.-grunat.02012003\_085613\_26c2 -RCFO=6 -RCFO=3  
-NO\_XLPHY -NO\_MAP -LANG=ENGLISH -NEG\_SCORES=0 -WAIT -LONGLIST -DEVTIMEOUT=12  
-WFRN-TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-WAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: PIR\_73.\*

1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	10.7	220	2	T40806
2	68	10.4	1040	2	A84505
3	66	10.0	155	2	A44890
4	66	10.0	349	2	A84954
5	65.5	10.0	431	2	T06019
6	65	9.9	310	2	T01266
7	64.5	9.8	269	1	R5BVL3
8	64.5	9.8	531	2	H90125
9	64	9.8	3712	1	YGC5VC
10	63.5	9.7	120	2	JE0177
11	63.5	9.7	532	2	B82354
12	63.5	9.7	760	2	G71417
13	63	9.6	305	2	A10336
14	63	9.6	365	2	JC7527

15	62.5	9.6	500	1	100115
16	62.5	9.6	690	2	100115
17	62.5	9.6	2100	2	100115
18	62	9.6	265	2	100115
19	62	9.6	326	2	100115
20	62	9.6	352	2	100115
21	62	9.6	437	1	100115
22	62	9.6	437	2	100115
23	62	9.6	1012	2	100115
24	62	9.6	1012	2	100115
25	62	9.6	1070	2	100115
26	62	9.6	1070	2	100115
27	61.5	9.4	1326	2	100115
28	61.5	9.4	1410	2	100115
29	61.5	9.4	1410	2	100115
30	61.5	9.4	455	2	100115
31	61	9.3	584	2	100115
32	61	9.3	739	2	100115
33	61	9.3	1035	2	100115
34	61	9.3	1035	2	100115
35	60.5	9.2	471	2	100115
36	60.5	9.2	471	2	100115
37	60.5	9.2	643	2	100115
38	60.5	9.2	905	2	100115
39	60	9.2	114	2	100115
40	60	9.2	584	2	100115
41	60	9.2	584	2	100115
42	60	9.2	1035	2	100115
43	59.5	9.1	140	1	100115
44	59.5	9.1	140	1	100115
45	59.5	9.1	265	2	100115

#### ALIGNMENTS

##### Result 1

T40806 Hypothetical protein SPB8B7.12c - fission yeast (Schizosaccharomyces pombe)

Citation: 03-Dec-1999 sequence\_revision 03-Dec-1999 \*text\_change 03-Dec-1999

Accession: T40806

RefSeq: A1 RefSeq: K1 Line: M1 RefSeq: M.A. Bartell, B. B.

Submitted to the EMBL Data Library: October 1998

Reference number: 21949

Accession: T40806

Status: preliminary translated from cDNA/EMBL/DBJ

Accession type: DNA

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

Accession: 100000000

fusca protein homo  
hypothetical prote  
transcription fact  
hypothetical prote  
hypothetical prote  
hypothetical prote  
conserved hypothet  
hypothetical prote  
membrane alpha pr  
glycine protein - fr  
zinc metalloprotei  
GTP cyclohydrolase  
probable WD-40 rep  
secy protein homol  
conserved hypothet  
zinc protein - Esc  
3-glycerin precursor  
AIP-dependent exod  
probable maturase  
5'-methylthioadeno  
translational initia  
fusca protein homo  
hypothetical prote  
preprotein translo  
G protein-coupled  
probable coenzyme  
eaf polyprotein -  
eaf polyprotein - s  
nucleophilic - fss  
conserved hypothet  
extragenic suppres

```

1b 157 lysylser-----AspTyrGlySerSerIleLeuIleValLeuAlaLeuLysIle 174
156 TCCATTCGCC-----ATACTGAGAAACTGCGAGAGAGCCAGTGTCTTACAGGTTCTG 246
155 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 175 AlaArgAlaIleArgTGTGTTLeuAlaLeuAlaArgAsnValIleGluIleLeuLeuLeu----- 191
174 111 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 247 AAGAAACTTCGATCATCTCTTCTGAGCGAAGAAACACCGAAGTCTTACATAAATAATCTTGG 305
146 111 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 192 ---GtSerGlyIleSerPheIleuGluAsnLysIleGly-----AlaIleTlleMetLysPhe 205
191 111 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 307 TGC 309
111
1b 210 Cys 210
111

RESULT 2
A:4505
Probable TNP2-like transposon protein (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
A:Accession: A84505
RefSeq: X: Kaul, S.; Rounsley, S.D.; Shen, L.P.; Beattie, M.L.; Town, S.D.;
Kellum, K.O.; Hoffman, K.S.; Croft, L.A.; Sheen, J.A.; Vankken, S.E.;
Hanus, D.; Niernann, W.C.; White, C.O.; Eisen, J.A.; Salzberg, S.L.;
Fischer, G.W.; Venter, J.
Nature 402, 761-766, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84505
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1040 <SNO>
A:Cross-References: GB:AE002093; MLD:q4432794; P1DN:AA020645.1;
C:Genetics:
A:Gene: At2g13000
A:Map position: 2
A:Map position: 2

Alignment Scores:
Pred. NO.: 3 56 Length: 1040
Score: 66.00 Matches: 15
Percent Similarity: 41.258 Conservative: 14
Best Local Similarity: 23.758 Mismatch: 29
Query Match: 10.358 Indels: 18
DB: 2 Gaps: 3

US-09-762-v27-4 (1-388) x A84505 (1-1040)
1b 153 TCAGTCGGTAGAGTATTAGCTACAGAAACCTTCTCATTCGCAATGAGAAAGTGGAG 222
152 111 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 350 SerMetMetThrIleLysAlaAspAsnAsnLeuSerGluLysCysMetAspSerIlePhe 359
349 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 223 CAGGCAATG-----TGGCTATGAGG 243
222 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 370 GluIleuIleLysGluIleuIleuProProAspAsnIleSerAlaLysSerIleGluIle 389
369 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 244 TACAACAAGAACTCAGATCATCTCTTTCAGGAGAAAGAACGACAG-----TGCATGAGAGAG 309
243 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 350 GluIleuLeuValSerSerHisGlyLeuProSerGluMetIleAspVal3CysIleAspTyr 419
349 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 301 TGCCTTGCTCTC-----ATACCTCCAGAAAGCTGAGATTCGATTAATAATCTT 348
300 111 111 111 111 111 111 111 111 111 111 111 111 111
1b 410 CysMetIlePheTrpGlyAspAspValAsnLeuGluIleCysArgPheG3SylIlePhe 429
409 111 111 111 111 111 111 111 111 111 111 111 111 111

RESULT 3
A:44850
E6 protein - human papillomavirus type 66
C:Species: human papillomavirus type 66
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 23-Mar-1993
A:Accession: A44850
RefSeq: X: Beaudenon, S.; Favre, M.; Orth, G.;
Ritwaleed, A.R.;
J. Clin. Microbiol. 29, 2656-2660, 1991
A:Title: Characterization of human papillomavirus type 66 from an invasive carcinoma of
A:Reference number: A44850; MUID:92129556; PMID:1663515
A:Accession: A44850

```

[illegible]





```

A:Accession: S27008
A:Molecule type: DNA
A:Residues: 1-138 "C",140-269 <GRA3>
A:Cross-references: EMBL:X65014; NID:q3563; PIDD:CA44614e.1; FID:q3563
A:Experimental source: strain 07173
R:Grack, H.R.; Grohmann, L.; Choll, T.
FEBS Lett. 242, 4-8, 1988
A>Title: Mitochondrial ribosomes of yeast: isolation of individual proteins and determination of their amino acid sequence
A:Reference number: S26754; MUID:85076818; PMID:3060376
A:Accession: S26756
A:Molecule type: protein
A:Residues: 21, V, 22-53 <GRA>
A:Experimental source: strain 07173
A:Note: This sequence has been revised in reference S23455
R:Grohmann, L.; Grack, H.R.; Kruf, V.; Choll, T.; Goldschmidt-Peisin, S.; Kitzawa, M
FEBS Lett. 264, 51-56, 1991
A>Title: Extended N-terminal sequencing of proteins of the large ribosomal subunit from yeast
A:Reference number: S17255; MUID:91285106; PMID:2060626
A:Accession: S26637
A:Molecule type: protein
A:Residues: 176, TP, 179-182, 257-261 <GRD>
R:van der Aart, Q.J.M.; Kline, K.; Steensma, H.Y.
submitted to the EMBL Data Library, June 1995
A:Description: Sequence analysis of the 43 KB CKM1-YLM9-PEI54-SM11-Pro-1-YMD4-PK11-1
A:Reference number: S57680
A:Accession: S57683
A:Molecule type: DNA
A:Residues: 1-269 <VAN>
A:Cross-references: EMBL:X87941; NID:q886508; PIDD:CA611e.1; FID:q886508
R:van der Aart, Q.J.M.; Steensma, H.Y.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64541
A:Accession: S64544
A:Molecule type: DNA
A:Residues: 1-269 <VAN>
A:Cross-references: EMBL:J73004; NID:q1323394; PIDD:CA4724e.1; FID:q1323394; SFFB:CA4724e.1
A:Experimental source: strain S288C
R:van der Aart, Q.J.M.; Kline, K.; Steensma, H.Y.
test 12, 385-390, 1996
A>Title: Sequence analysis of the 43 KB CKM1-YLM9-PEI54-DIE2-SM11-Pro-1-YMD4-PK11-1
A:Reference number: S63896; MUID:96267763; PMID:8701610
A:Accession: S63899
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-269 <VAN>
A:Cross-references: EMBL:X87941; NID:q886508; PIDD:CA611e.1; FID:q886508
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1995
C:Genetics:
A:Gene: SGD:MRP19; YML5; MIPS:YGR220C
A:Cross-references: SGD:S0003452; MIPS:YGR220C
A:Map position: 7R
A:Genome: nuclear
C:Superfamily: Escherichia coli ribosomal protein L3
C:Keywords: mitochondrion; protein biosynthesis; ribosome
F.1-19/Domains: transit peptide (mitochondrion) *status experimental *S172
F.20-269/Product: ribosomal protein L3, mitochondrial *status experimental *S267

```

[illegible]

[illegible][illegible]





Genome version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: January 2, 2003, 14:01:48 : Search time 15 seconds  
(without alignments)  
2475.818 Million cell updates/sec

File: US-09-762-027-4

Perfect score: 657  
Sequence: 1 agctgcatacagacat.....atccataaangtutanger jee

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delpext 7.0	

Searched: 112892 seqs, 4176328 residues

Total number of hits satisfying chosen parameters: 25744

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus.model -DEV=xlh  
-C=/C:/usr/local/ncbi/blast/blastn -FASTA=us-09-762-027-4.n2p.rsp -MINMATCH=0.1 -MAXPCT=0  
-DB=SwissProt\_10 -OPMT=fastan -SUFFIX=02p -MNAME=us-09-762-027-4.n2p.rsp -MFILE=us-09-762-027-4.n2p.rsp  
-LOOKUP=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -HITS=human40.git  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=PCT -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09762027 -FCGN\_1\_1\_6 -runat\_02012003\_085613 -2642 -NCI=6 -ICR043  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGJOB -EV -ITER=2048 -15  
-NARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=0 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: SwissProt\_10\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67.5	10.3	807	1	YNG1_SCHPO
2	66	10.0	155	1	VE6_HPV66
3	66	10.0	349	1	GUAC_BUCAT
4	64.5	9.8	269	1	RM09_YEAST
5	64	9.8	215	1	US20_CAHEL
6	63.5	9.7	371	1	ACV5_CBPAC
7	63.5	9.7	120	1	SV16_HUMAN
8	62.5	9.5	471	1	CSN1_HUMAN
9	62.5	9.5	1400	1	RIFI_SCHPO
10	62	9.4	437	1	Y269_HELPY
11	62	9.4	437	1	Y269_HELPY
12	61.5	9.4	192	1	GC82_HELPY
13	61	9.3	286	1	VNST_INCYA
14	61	9.3	455	1	YHFN_ECOLI
15	61	9.3	599	1	HM21_HUMAN
16	61	9.3	1564	1	N181_SCHPO
17	60.5	9.2	164	1	CGHB_CALJA
18	60.5	9.2	252	1	Y060_METJA

19	60.5	9.2	252	1	Y060_METJA
20	60.5	9.2	252	1	Y060_METJA
21	60.5	9.2	252	1	Y060_METJA
22	60.5	9.2	252	1	Y060_METJA
23	60.5	9.2	252	1	Y060_METJA
24	60.5	9.2	252	1	Y060_METJA
25	60.5	9.2	252	1	Y060_METJA
26	60.5	9.2	252	1	Y060_METJA
27	60.5	9.2	252	1	Y060_METJA
28	60.5	9.2	252	1	Y060_METJA
29	60.5	9.2	252	1	Y060_METJA
30	60.5	9.2	252	1	Y060_METJA
31	60.5	9.2	252	1	Y060_METJA
32	60.5	9.2	252	1	Y060_METJA
33	60.5	9.2	252	1	Y060_METJA
34	60.5	9.2	252	1	Y060_METJA
35	60.5	9.2	252	1	Y060_METJA
36	60.5	9.2	252	1	Y060_METJA
37	60.5	9.2	252	1	Y060_METJA
38	60.5	9.2	252	1	Y060_METJA
39	60.5	9.2	252	1	Y060_METJA
40	60.5	9.2	252	1	Y060_METJA
41	60.5	9.2	252	1	Y060_METJA
42	60.5	9.2	252	1	Y060_METJA
43	60.5	9.2	252	1	Y060_METJA
44	60.5	9.2	252	1	Y060_METJA
45	60.5	9.2	252	1	Y060_METJA

# ALIGNMENTS

## RESULT 1

YNG1\_SCHPO 2475.818 607 AA

AC 01-NOV-1998 (F00009) (Uncloned)

D1 15-DEC-1998 (F00037) (last sequence update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)

D1 15-DEC-1998 (F00040) (last annotation update)







[illegible][illegible]



```

SQ SEQUENCE 120 AA: 13600 MW: 373073016134094 CRGc4:
Alignment Scores:
Pred. NO.: 5-27 Length: 120
Score: 63.50 Matches: 16
Percent Similarity: 47.928 Conserved: 7
Best Local Similarity: 33.338 Mismatches: 22
Query Match: 9.67% Indels: 3
DB: 1 Gaps: 2
US-QY-761-Q27-4 (1-368) x SY16_HUMAN (1-120)
QY 214 AACGGCAGACAGCAAGCTGCCAACAGGCTAC--AAAGAAATTCATCATCATCGTCTTT
||| |||||||::: ||| ||| ||| ||| ||| |||
Cb AsnThrProSerThrcysLeuAlaLysTrpArgValcysValleProAspArgLeval 51
QY 221 GAAGGA-----AAGAAGCTGAGGTCTACATTAAGACTCTTGCTGCATACCTCTGGA 324
||| ||| ||| |||::: ||| ||| ||| ||| ||| |||
Db ValGIyTYrGLyALeUasncysHisLeuProAlaIlellePheValThrLysArg 71
QY 325 AAGTCGACATTCTGTATAATGCT 348
::: ||| |||||
Db 72 AsnArgGLuValcysThrAsnPro 79
RESULT 8
CSN1_HUMAN STANDARD PRI: 471 AA.
ID CSN1_HUMAN Q1056; OSBML1;
AC Q1056; OSBML1; (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
CD COP1 signalosome complex subunit 1 (G protein pathway suppressor 1)
DE (Gpsi protein) (MFW protein).
GN GPST OR COPSL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
ON NCBI_TaxId=9606;
RN [1]
RP MEDLINE=97098647; PubMed=8543324;
RX Chan B.H., Bowdsh K.S., Pacal A., Fluckiger Staub S., Koos L.,
RA Spahn R.Y., Xie W., Colicelli J.;
RA "Two human CDAs, including a homolog of Arabidopsis FUS3 (Gpsi1),
RI suppress G-protein- and mitogen-activated protein kinase-mediated
RT signal transduction in yeast and mammalian cells.";
RL Mol. Cell. Biol. 16:6656-6706(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: SUPPRESSES G-PROTEIN- AND MITOGEN-ACTIVATED PROTEIN
KINASE-MEDIATED SIGNAL TRANSDUCTION.
CC -! SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -! SIMILARITY: CONTAINS 1 PCI DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Collection.
CC The European Bioinformatics Institute. There are no restrictions on its use
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib.slb.ch).
CC -----
DR EMBL: U20285; AAC0906.2; ALT_INTF.
DR EMBL: BC000155; AAH00155.1; -.
DR Genbank: HGNC:4549; GPST.
DR MIM: 601934; -.
DR InterPro: IPR000717; PCl.
DR Pfam: PF01359; PCl_1.
DR SMART: SM00086; PINTf_1.
RW Signalosome: Nuclear protein.

```

[illegible]





**INFLUENZAVIRUS C**  
GenBank  
Seq\_ID=115087  
Accession: F62911177  
Source: PubMed-2943076  
Author: Nakada S., Fitch W.H., Palese P.  
Title: Influenza C virus in man: multiple evolutionary lineages and low rate of change.\*  
Journal: Virology 153:12-21(1986).

This SNUS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL database at European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/about-us> or send an email to [license@isb.sib.ch](mailto:license@isb.sib.ch)).

EMBL D00033; BAA24042.1;  
EMBL D00033; BAA24043.1;  
GenePro: IPRO05187; Flu\_C.NSI.  
InterPro: IPRO05188; Flu\_C.NS2.  
Pfam: PF03506; Flu\_C.NSI; 1.  
Pfam: PF03505; Flu\_C.NS2; 1.  
Structural protein.  
Chain 1 286 NS1 PROTEIN.  
Chain 1 62 NS2 PROTEIN.  
SEQUENCE 260 AA; 32436 MW; EABD5BDB8C897FBD5 CRC64;

Alignment Scores:  
Pred. Score: 11.5 Length: 266  
Percent Similarity: 61.00 Matches: 21  
Best Local Similarity: 38.55% Conservative: 11  
Query Match: 25.30% Mismatch: 36  
Gap: 9.28% Indels: 15  
Gaps: 1

OS-09-02-027-4 (1-388) x VNST\_INCYA (1-286)

45 AAGCGCTGATTTCTTGTCATGGACAGACGTGGTTTAATCACTACGAAGAACAATGCACGATA 1-4  
||||| ||| :||| |::| ::| :||| ||| ||| |||  
1 Lysihcyssyltyrleutllealaarglnvalalaglyile-cluhirgylleargy 1-6  
15 CTTCGGAACCTTFATGCTTACAGTACACGACGACGACGACGACGACGACGACGACGACGACGAC 1-4  
111 ||| |||  
10 scys-----Pheargysilidy 1-6  
15 AGTCGCCGATATTACGTAACAGAACACTTCCATTCGCACACTACGAAGAACAATGCACGATA 2-4  
||||| ||| :||| |::| ::| :||| ||| ||| |||  
15 sserglyphevalmetalelthaspolltleserleuthrltleunaryst lleyssz 2-6  
--5 GGCACGTGTCCTACAGGCTACAGAACAACTTCAGATCATCTCTTGAGGAGAAAGACT 2-4  
||||| ||| ||| |||  
16 rglvalagltnleaspprolyrtfrrpolmuretyshnglnleleuthra;pleatyle 2-6  
--5 GAAGTGC 2-1  
|||||  
16 -ro ulleycs 2-6

RESULTS  
YMFN\_ECOLI STANDARD; PRT; 455 AA.

BT 1-JUL-1998 (Rel. 36, Last Created)  
BT 1-JUL-1998 (Rel. 36, Last sequence update)  
BT 1-OCT-2001 (Rel. 40, Last annotation update)  
RE Biotechnological protein ymfN.  
RE YMFN OR B1145.  
SS Escherichia coli.  
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
CC Escherichia.  
CC Refl\_TextID=562;  
CC 1-1

1. **SEQUENCE FROM N.A.**  
 2. **STRAIN K-12**  
 3. **REFERENCE**  
 4. **REFERENCE**  
 5. **REFERENCE**  
 6. **REFERENCE**  
 7. **REFERENCE**  
 8. **REFERENCE**  
 9. **REFERENCE**  
 10. **REFERENCE**  
 11. **REFERENCE**  
 12. **REFERENCE**  
 13. **REFERENCE**  
 14. **REFERENCE**  
 15. **REFERENCE**  
 16. **REFERENCE**  
 17. **REFERENCE**  
 18. **REFERENCE**  
 19. **REFERENCE**  
 20. **REFERENCE**  
 21. **REFERENCE**  
 22. **REFERENCE**  
 23. **REFERENCE**  
 24. **REFERENCE**  
 25. **REFERENCE**  
 26. **REFERENCE**  
 27. **REFERENCE**  
 28. **REFERENCE**  
 29. **REFERENCE**  
 30. **REFERENCE**  
 31. **REFERENCE**  
 32. **REFERENCE**  
 33. **REFERENCE**  
 34. **REFERENCE**  
 35. **REFERENCE**  
 36. **REFERENCE**  
 37. **REFERENCE**  
 38. **REFERENCE**  
 39. **REFERENCE**  
 40. **REFERENCE**  
 41. **REFERENCE**  
 42. **REFERENCE**  
 43. **REFERENCE**  
 44. **REFERENCE**  
 45. **REFERENCE**  
 46. **REFERENCE**  
 47. **REFERENCE**  
 48. **REFERENCE**  
 49. **REFERENCE**  
 50. **REFERENCE**  
 51. **REFERENCE**  
 52. **REFERENCE**  
 53. **REFERENCE**  
 54. **REFERENCE**  
 55. **REFERENCE**  
 56. **REFERENCE**  
 57. **REFERENCE**  
 58. **REFERENCE**  
 59. **REFERENCE**  
 60. **REFERENCE**  
 61. **REFERENCE**  
 62. **REFERENCE**  
 63. **REFERENCE**  
 64. **REFERENCE**  
 65. **REFERENCE**  
 66. **REFERENCE**  
 67. **REFERENCE**  
 68. **REFERENCE**  
 69. **REFERENCE**  
 70. **REFERENCE**  
 71. **REFERENCE**  
 72. **REFERENCE**  
 73. **REFERENCE**  
 74. **REFERENCE**  
 75. **REFERENCE**  
 76. **REFERENCE**  
 77. **REFERENCE**  
 78. **REFERENCE**  
 79. **REFERENCE**  
 80. **REFERENCE**  
 81. **REFERENCE**  
 82. **REFERENCE**  
 83. **REFERENCE**  
 84. **REFERENCE**  
 85. **REFERENCE**  
 86. **REFERENCE**  
 87. **REFERENCE**  
 88. **REFERENCE**  
 89. **REFERENCE**  
 90. **REFERENCE**  
 91. **REFERENCE**  
 92. **REFERENCE**  
 93. **REFERENCE**  
 94. **REFERENCE**  
 95. **REFERENCE**  
 96. **REFERENCE**  
 97. **REFERENCE**  
 98. **REFERENCE**  
 99. **REFERENCE**  
 100. **REFERENCE**

```

14 OCT 12
295 Gln 299

RESULT 15
HM21_HUMAN STANDARD: PRT: 595 AA.
ID 090605: 090605: 075673: 075672:
AC 16-OCT-2001 (Rel. 40. Created)
DT 15-OCT-2001 (Rel. 40. Last sequence update)
DI 15-JUN-2002 (Rel. 41. Last annotation update)
DE High-mobility group protein 2-like 1 (HMGBC protein).
UN HMG21L1 OR HMGBCG.
CS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
CC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
KA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
KA Clump M., Smink L.J., Altschough R., Almeida J.P., Babbage A.R.,
KA Bagley C., Bailey J., Barlow K.F., Bates K.N., Beasley J.P.,
KA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
KA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
KA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor K.,
KA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
KA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.S.,
KA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
KA Gilbert J.G.R., Goward M.E., Graffam D.V., Griffiths M.N.D., Hall C.,
KA Hall R.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,
KA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., Lloyd D.M.,
KA Laid G.R., Langford G.F., Leversha M.A., Lloyd G., Lloyd D.M.,
KA Martin I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
KA McCall J., McLaren S., McNair A.A., Milne S.A., Mortimore B.J.,
KA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.I.,
KA Phillips S.H., Plumb R.W., Ramsey H., Ramsey L., Ross M.L.,
KA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
KA Soderlund C., Spragon L., Stewart C.A., Sulston J.E., Swann R.M.,
KA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Wiley D.L.,
KA Williams L., Williams S.A., Williamson H., Wilmer I.E., Wilming L.,
KA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
KA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
KA Shimada A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
KA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do I.,
KA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.L.,
KA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen L., Pan H.,
KA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaili S., Sloan D., Song L.,
KA Wang Q., Wang Y., Wang Z., White J., Williamson D., Wu H., Yao Z.,
KA Zhan M., Zhang G., Chisose S., Murray J., Miller N., Mink P.,
KA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw B., Bourne S.,
KA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
KA Hinds K., Kemp K., Lattelle P., Layman D., Ozeresky P., Roffing I.,
KA Scher P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
KA Kort I., Bedell J.A., Hillier L., Mardis E., Waterston K., Wilson R.,
KA Emanuel B.S., Shaikh T., Kurahashi H., Saito S., Budarf M.L.,
KA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelman L.,
KA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard F., Kieda D.,
KA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
KA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
KA Tikhonov Y., Wright H.;
RA *The DNA sequence of human chromosome 22.*;
RI Nature 402:489-495(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX Collins J.E., Huckle E.J.;
RI Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-390 AND 406-559 FROM N.A.
RX MEDLINE=95263500; PubMed=10329004;
KA Seroussi E., Kieda D., Kost-Alimova M., Sandberg-Nordqvist A.,
KA Fransson I., Jacobs J.F., Fu Y., Pan H.-Q., Roe B.A., Imreh S.,

```





XX 08-FEB-2001: 2001MO-US04423.  
 XX  
 XX 11-FEB-2000: 2000US-0181856.  
 XX 17-FEB-2000: 2000US-0183684.  
 XX 25-FEB-2000: 2000US-0185141.  
 XX 03-MAR-2000: 2000US-0186818.  
 XX 05-MAR-2000: 2000US-0188345.  
 XX 17-MAR-2000: 2000US-0189597.  
 XX  
 XX (INCYTE) INCYTE GENOMICS INC.  
 XX  
 XX Tang YT, Yue H, Baughn MR, Yao MG, Bandman O, Azimzai Y, Tai P,  
 XX Gauthi AR, Ring HZ, Shih LL, Yang J, Pollocky JL:  
 XX WPI: 2001-514673/56.  
 XX N-PSDB: AAH76158.  
 XX  
 XX Isolated polypeptide encoding a drug metabolizing enzyme useful for the  
 XX diagnosis, treatment, and prevention of autoimmune/inflammatory, cell  
 XX proliferative, developmental and endocrine disorders.  
 XX  
 XX Claim 1: Page 130: 150pp: English.  
 XX  
 XX The invention provides human drug metabolizing enzymes (DME) and  
 XX polynucleotides encoding the DMEs. The DME can be expressed by standard  
 XX recombinant methodology. DMEs and their agonists and antagonists are  
 XX useful for the diagnosis, treatment, and prevention of autoimmune/  
 XX inflammatory, cell proliferative, developmental, endocrine such as  
 XX aneurysm, eye, metabolic, and gastrointestinal disorders, including liver  
 XX disorders and infection. The present sequence represents a human DME.  
 XX  
 XX Sequence 294 AA:  
 XX  
 XX Alignment Scores:  
 XX Fred. No.: 2:25e-07 Length: 294  
 XX Score: 117.00 Matches: 27  
 XX Percent Similarity: 53.10% Conservative: 0  
 XX Best Local Similarity: 53.10% Mismatches: 1  
 XX Query Match: 17.81% Indels: 1  
 XX Gaps: 0  
 XX  
 XX US-09-762-027-4 (1-368) x AAB85771 (1-294)  
 XX  
 XX 300 AIGCTTGTCTTCAATCACTCAGAGCTCCGATTCGTATTAATCCTTASAAAGAGAG 359  
 XX 1 MetLeuValLeuHisAsnSerGlnIleLeuGlnIleLeuTyrTysSerLeuGlnIleLeuTyr 29  
 XX  
 XX 360 ATCCCGTCGATCCATTAANGATATAT 384  
 XX 21 Ile-ProIleSerIleLeuValTyr 28  
 XX  
 XX RESULT 2  
 XX AAB56487 standard: Protein: 241 AA.  
 XX  
 XX AAB56487:  
 XX  
 XX 13-MAR-2001 (first entry)  
 XX  
 XX Human prostate cancer antigen protein sequence SEQ ID No:1905.  
 XX  
 XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
 XX neutroprotective; cytoskeletal; cardiovascular; immunomodulatory; muscular;  
 XX vulnereary; gastrointestinal; nephrotoxic; antineoplastic; gynaecological;  
 XX antibacterial; gene therapy; neural; immune; reproductive; renal;  
 XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
 XX wound; infectious disease.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO200055174-A1.

XX 21-Sep-2000:  
 XX  
 XX 08-MAR-2000: 2000US-0805928.  
 XX  
 XX 12-MAR-1999: 99-5-0124279.  
 XX  
 XX (HUMAN) HUMAN GENOME SOLI 167.  
 XX (ROSE/) ROSEN G A  
 XX  
 XX Rosen GA, Kucher JO:  
 XX WPI: 2000-58713/75.  
 XX N-PSDB: AAF15700.  
 XX  
 XX Prostate cancer associated gene sequences, referred to as prostate  
 XX cancer antigens, useful for treatment, prevention, and diagnosis of  
 XX disorders such as prostate cancer.  
 XX  
 XX Claim 1: Page 142-148; 236pp: English.  
 XX  
 XX AAF1566 to AAF1575 encode the human prostate cancer associated  
 XX proteins, called prostate cancer antigens, given in AAB5363 to AAB57302.  
 XX The prostate cancer antigens can have neutroprotective, cytoskeletal,  
 XX cardiovascular, immunomodulatory, muscular, vulnereary, gastrointestinal,  
 XX nephrotoxic, antineoplastic, gynaecological and antibacterial activities,  
 XX and can be used in gene therapy. The prostate cancer antigen  
 XX polynucleotides may be used for detection of prostate cancer, circumsome  
 XX identification, as chromosome markers, and for numerous other diagnostic  
 XX or research purposes. The prostate cancer antigens may be used to treat  
 XX disorders such as neural, immune, muscular, reproductive,  
 XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 XX disorders, wounds and infectious diseases. AAF1566 to AAF1574 to  
 XX AAB5703 represent sequences used in the exemplification of the present  
 XX invention.  
 XX  
 XX Sequence 241 AA:  
 XX  
 XX Alignment Scores:  
 XX Fred. No.: 0.00104 Length: 241  
 XX Score: 91.00 Matches: 20  
 XX Percent Similarity: 77.14% Conservative: 7  
 XX Best Local Similarity: 57.14% Mismatches: 7  
 XX Query Match: 13.85% Indels: 1  
 XX Gaps: 0  
 XX  
 XX US-09-762-027-4 (1-368) x AAB56487 (1-241)  
 XX  
 XX 28 TTTAAGAGTATATTAAGATGCTTGGCTTGTATTAATCCTTASAAAGAGAG 341  
 XX 9 SerIleAlaSerSerArgMetIleLeuLeuAsnSerHisTyrIleLeuValTyr 28  
 XX  
 XX 342 AAATGTTTAAATAGAGATGCTTGGCTTGTATTAATCCTTASAAAGATAT 384  
 XX 29 LysSerLeuAlaArgLysSerIle-ProIleSerIleLeuValTyr 42  
 XX  
 XX RESULT 2  
 XX AAB56575 standard: Protein: 190 AA.  
 XX  
 XX AAB56575:  
 XX  
 XX 21-JUN-2001 (first entry)  
 XX  
 XX Human; chronic leukocytic leukemia;  
 XX Chronic leukocytic leukemia; cancer; chemotherapy; leukemia; gastritis;  
 XX tumor; cell mediated autoimmune disease; liver cirrhosis;  
 XX osteoarthritis; fibromyofibrosis; schistosomiasis; trichinosis;  
 XX ascariasis; prostaglandin-independent fever; siliosis;  
 XX rheumatoid arthritis; hyper-eosinophilic syndrome.  
 XX  
 XX Homo sapiens.  
 XX  
 XX

```

Key Location/Qualifiers
Peptide Loc_16
Label= Signal_peptide
Protein /label= chemokine_beta-1-
IN W096JY520-A1.
ID 12-FEB-1996.
XN uc-JUN-1995; 95MO-USO7171.
XX uc-JUN-1995; 95MO-USO7171.
XN uc-JUN-1995; 95MO-USO7171.
FA (HMA-) HUMAN GENOME SCI INC.
FA (SMR) SMTIHLTRLE BEECHAM CORP.
XN 11 F. Mooney JL.
XR WP: 1997-043142/04.
XR M-CSDB; AAI49221.
XX Isolated human chemokine beta-12 and corresponding cDNA sequence used to develop probes for treating e.g., tumours, auto-immune diseases and inflammatory and infective diseases
F1 Claim 1; Page 43; 55pp; English.
XX The present sequence represents the polypeptide comprising amino acids -16 to 120 chemokine beta-12 (CXbeta-12). CXbeta-12 can be used for the diagnosis and treatment of conditions associated with expression of the polypeptide. In particular, it can be used to inhibit the proliferation of haematopoietic cells such as bone marrow stem cells. It can be used for adjunct protective treatment during cancer chemotherapy or to treat psoriasis, to treat solid tumours, to enhance host defences against resistant chronic and acute infections, to treat T-cell mediated autoimmune diseases, to stimulate wound healing and prevent scarring; to treat fibrotic disorders such as liver cirrhosis, cystic fibrosis and pulmonary fibrosis, to regulate haemtopoiesis, to treat rheumatoid arthritis, osteoarthritis or spondylitis, or to treat cancers. CXbeta-12 analogues can be used e.g. in the treatment of certain autoimmune diseases like multiple sclerosis, rheumatoid arthritis, Crohn's disease, ulcerative colitis, Sjogren's syndrome, myasthenia gravis, systemic lupus erythematosus, idiopathic thrombocytopenic purpura, pernicious anemia, insulin dependent diabetes mellitus, Behçet's disease, Kawasaki disease, aplastic anaemia, sickle cell disease, eosinophilic syndrome.
```

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```
CC occurring genomic sequences. HIFs, and its catalytic or immunoregulatory CC fragments are useful for screening libraries of compounds in a variety CC drug screening techniques. Antibodies which specifically bind HIF may CC used for the diagnosis of disorders associated with overexpression of CC HIFs, or in assays to monitor patients being treated with HIF-α CC agonists, antagonists or inhibitors of HIFs. The present sequence CC represents an HIFs protein of the invention.
```

XX

Sequence      302 AA:

SQ

Alignment Scores:

Pred. No.:		Length:	Juz
Score:	77.00	Matches:	16
Percent Similarity:	79.31%	Conservative:	5
Best Local Similarity:	62.07%	Mismatches:	5
Query Match:	11.72%	Indels:	1
DG:	22	Gaps:	0

US-09-762-027-4 (1-388) x AAB73512 (1-302)

CY    300 ATGGTCGTGCCTCATACACTCAGCAAGTCCAGATTCGTAATAAACCCTTAGAAAACAAC 352  
             | | | | | | | | | | | | | | | | | | | | |  
CU     1 MetleuleubasAsarSerHisIshenleualaleunlrylsSertioAlalaRaser 20  
  
CY    360 ATCCCCGAATCGATTAAANGTATAT 364  
             ||| ||||| : :::::  
CU     21 lle-ProgluserteulsValTyI 26  
  
RESULT 5  
AABBY7238  
ID    ABBY7238 standard; Protein: 302 AA.  
XX  
XX    ABBY7238; N  
DI    27-JUN-2002 (first entry)  
XX  
DE    Novel human protein SEQ ID NO: 506.  
XX  
KM    Human; anti-nauseic; vulnerary; antinflammatory; immunomodulator;  
KM    antidepressant; cerebroprotective; cytoskeletal; rheumatic; gene therapy;  
KM    neuroprotective; anticancerous; protein therapy; ESI:  
XM    expressed sequence tag.  
XX  
JM    Homo sapiens.  
JN    WO200222660-A2.  
XX  
PD    21-MAR-2002.  
XX  
PF    10-SEP-2001; 2001WO-US36015.  
XX  
PK    11-SEP-2000; 2000US-0659671.  
PA  
PZ    (HYSE-) HYSEQ INC.  
PI    Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QH, Fan F,  
PI    Xue AJ, Yang Y, Weinman T, Brumac R:  
DR    WPI: 2002-292408/33.  
DR    N-PDB: ABN32424.  
PT    An isolated polynucleotide for treating diseases associated with its  
PS    encoded polypeptide such as cancer and multiple sclerosis .  
XS  
XS    Example 2; SEQ ID NO 506; 509pp; English.  
CS  
OC    The present invention provides the protein and coding sequences of the  
OC    novel human proteins. These were isolated from expressed sequences tags  
OC    (ESTs). They can be used to stimulate cell growth, to regulate  
OC    haematopoiesis e.g., to treat aplastic anaemia, to help tissue repair;  
OC    e.g. in burn treatment, to regulate the immune system e.g. to treat  
OC    multiple sclerosis, to regulate activity or inhibit e.g. to treat  
OC    infertility, to regulate haemostasis or thrombolysis e.g. to treat

Seq	Selected	Alignment Start	Score	Length	Matches	Mismatches	Indels	Gaps
1		1	100	302	16	5	1	0
2		1	100	302	16	5	1	0
3		1	100	302	16	5	1	0
4		1	100	302	16	5	1	0
5		1	100	302	16	5	1	0
6		1	100	302	16	5	1	0
7		1	100	302	16	5	1	0
8		1	100	302	16	5	1	0
9		1	100	302	16	5	1	0
10		1	100	302	16	5	1	0
11		1	100	302	16	5	1	0
12		1	100	302	16	5	1	0
13		1	100	302	16	5	1	0
14		1	100	302	16	5	1	0
15		1	100	302	16	5	1	0
16		1	100	302	16	5	1	0
17		1	100	302	16	5	1	0
18		1	100	302	16	5	1	0
19		1	100	302	16	5	1	0
20		1	100	302	16	5	1	0
21		1	100	302	16	5	1	0
22		1	100	302	16	5	1	0
23		1	100	302	16	5	1	0
24		1	100	302	16	5	1	0
25		1	100	302	16	5	1	0
26		1	100	302	16	5	1	0
27		1	100	302	16	5	1	0
28		1	100	302	16	5	1	0
29		1	100	302	16	5	1	0
30		1	100	302	16	5	1	0
31		1	100	302	16	5	1	0
32		1	100	302	16	5	1	0
33		1	100	302	16	5	1	0
34		1	100	302	16	5	1	0
35		1	100	302	16	5	1	0
36		1	100	302	16	5	1	0
37		1	100	302	16	5	1	0
38		1	100	302	16	5	1	0
39		1	100	302	16	5	1	0
40		1	100	302	16	5	1	0
41		1	100	302	16	5	1	0
42		1	100	302	16	5	1	0
43		1	100	302	16	5	1	0
44		1	100	302	16	5	1	0
45		1	100	302	16	5	1	0
46		1	100	302	16	5	1	0
47		1	100	302	16	5	1	0
48		1	100	302	16	5	1	0
49		1	100	302	16	5	1	0
50		1	100	302	16	5	1	0
51		1	100	302	16	5	1	0
52		1	100	302	16	5	1	0
53		1	100	302	16	5	1	0
54		1	100	302	16	5	1	0
55		1</						



```

FR 23-MAR-2000: 2000US-191637P.
FR 11-JUL-2000: 2000US-061415U.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW:
XX
XX WPI: 2001-656860/75.
XX
XX N-PSDB: ABL07759.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosures: SEQ ID NO 17760; 21pp + Sequence Listing: English.
XX
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention
XX is useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins.
XX (ABB57737-ABB72072).
XX
XX The sequence data for this patent did not form part of the prior art
XX specification, but was obtained in electronic format directly from WIP
XX at ftp.wipo.int/pub/published_pat_sequences.
XX
XX
XX Sequence 1229 AA:
XX
XX
XX Alignment Scores:
XX Pred. No.: 2.84 Length: 1229
XX Score: 68.50 Matches: 22
XX Percent Similarity: 43.86% Conservative: 26
XX Best Local Similarity: 19.10% Mismatches: 39
XX Query Match: 10.43% Indels: 25
XX DB: 22 Gaps: 4
XX
XX US-05-762-027-4 (1-388) x ABB63656 (1-1229)
XX
XX QY 13 CGGAACATATCTCAGTCCTCTCTGCTACAAAGCCCTGAGATTCCTGCTACAGAA 72
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 122 Arganillepe-----LeulPhealaleuProltripleLeulthPheVal----- 137
XX
XX QY 73 GAACCTCTGTTTAATCTACAGAACCCAAACTTCCTGAGAGCTTAAGTCTAACAGTA 132
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 138 -----AlserGlyLeuIIleLeuIIleLeuIIleLeuIIleLeuIIleLeuIIle 143
XX
XX QY 133 CAACGTGAGTCAGTGAATTTATGTTTCAGTTCGTAGAAATTAAT----- 153
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 150 -----ValIIleGlnHisPheCysIIleSerAsnTrpArgAlaIIleGlnLeuGly 154
XX
XX QY 184 -----AACGAACCTTCCATTCCTCCCTACAGAAACCTGAGG 162
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 167 GluLeuGlnGlyLeuPheTrpHisArgSerIleHisValIleLeuIIlePheGlyTrpSer 167
XX
XX QY 223 CAGCGAGTCGTCCCTACAGGCTCTACAAAGAACTTCAGATCACTTCAGGAGAAAGAA 262
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 187 GlnIIleIIleGlyGlyValSerLeuLeuLysAlaSerHisLeuSerGlyAsnSerArg 265
XX
XX QY 283 CTGAACTGCTACATAGAGTCTGTCTCTCTCACTACTCTGGA 324
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 207 ValTyrSerAlaValAsnIIleIIleIIleLeuTyrThrMetArg 220
XX
XX
XX RESULT 5
XX ABB62231
XX ID ABB62231 standard; Protein: 2441 AA.
XX
XX AC ABB62231:
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 1-1405.

```

[illegible]





















[illegible][illegible]





Fri Jan 3 14:27:00 2003

us-09-762-027-4.n2p.rspt

Page 8

Best Local Similarity:	27.598	Mismatches:	6
Query Match:	9.89%	Indels:	18
DB:	10	Gaps:	4

```
Query Match: 9.898
DB: 10
```

10

Caps:

4-2

4-2

US-09-762-027-4 (1-388) x 064924 (1-310)

[illegible]

```

search completed: January  2, 2003, 15:11:07
Job time : 49.5 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Computer Ltd.

CM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 2, 2003, 15:09:29 : Search time 13.5 seconds

(without alignments)  
1691.273 Million cell updates/sec

Title: US-09-762-027-4

Perfect score: 657

Sequence: 1 agctgcacatacagcaacat.....atccataatgataatgag jss

Scoring table:

BLIUSOM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Rgapop 6.0 , Rgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 26422922 residues

Total number of hits satisfying chosen parameters: 525146

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_n2p model -DEV=xlh  
-O=/cgm2\_1/USP/OUS09762027/rumac\_02012003\_085614\_2671/app\_query\_fas\_1.1.553  
-DB=Issued\_Patents\_AA -QFMT=faslan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -L=ACPGT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blusom62 -TRANS=human45.cth  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=500 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=20000000  
-USER=0509762027 @CGN1.1.4 jrunal.02012003.085614.2671 -NCPU=6 -ICPU=3  
-NO\_XMAP=NO\_XMAP -LARGOUDRY -NEG\_SCORES=0 -WAIT -LONGLOC -DEV\_TIMEOUT=15  
-MARK\_TIMEOUT=30 -THEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -YGAPOP=6 -YGAPEXT=7  
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_AA:  
1: /cgm2\_6/ptodata/2/laa/5A.COMB.pep.\*  
2: /cgm2\_6/ptodata/2/laa/5B.COMB.pep.\*  
3: /cgm2\_6/ptodata/2/laa/6A.COMB.pep.\*  
4: /cgm2\_6/ptodata/2/laa/6B.COMB.pep.\*  
5: /cgm2\_6/ptodata/2/laa/PTUS.COMB.pep.\*  
6: /cgm2\_6/ptodata/2/laa/backfilltest.pep.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being plotted;  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	12.6	136	5	PCT-US95-07171-2
2	83	12.6	191	5	PCT-US95-07171-3
3	83	9.7	892	1	US-07-977-434-12
4	64	9.7	892	1	US-08-158-819-12
5	64	9.7	892	5	PCT-US91-07035-14
6	64	9.8	3665	2	US-08-222-617A-1
7	64	9.8	3712	2	US-08-222-617A-4
8	64	9.8	3712	2	US-08-222-617A-25
9	64	9.7	120	1	US-08-347-462B-2
10	63.5	9.7	120	2	US-08-798-143-2
11	63.5	9.7	120	5	PCT-US95-15484-2
12	61	9.3	246	2	US-08-730-258-2

13	61	9.3	412	1	US-09-447-540A-8
14	61	9.3	412	1	US-09-447-540A-9
15	59	9.3	114	1	US-06-129-704-2
16	59	9.3	114	1	US-06-129-704-2
17	59	9.3	114	1	US-06-129-704-2
18	59	9.3	114	1	US-06-129-704-2
19	59	9.3	114	1	US-06-129-704-2
20	59	9.3	114	1	US-06-129-704-2
21	59	9.3	114	1	US-06-129-704-2
22	59	9.3	114	1	US-06-129-704-2
23	59	9.3	114	1	US-06-129-704-2
24	59	9.3	114	1	US-06-129-704-2
25	59	9.3	114	1	US-06-129-704-2
26	59	9.3	114	1	US-06-129-704-2
27	59	9.3	114	1	US-06-129-704-2
28	59	9.3	114	1	US-06-129-704-2
29	59	9.3	114	1	US-06-129-704-2
30	59	9.3	114	1	US-06-129-704-2
31	59	9.3	114	1	US-06-129-704-2
32	59	9.3	114	1	US-06-129-704-2
33	59	9.3	114	1	US-06-129-704-2
34	59	9.3	114	1	US-06-129-704-2
35	59	9.3	114	1	US-06-129-704-2
36	59	9.3	114	1	US-06-129-704-2
37	59	9.3	114	1	US-06-129-704-2
38	59	9.3	114	1	US-06-129-704-2
39	59	9.3	114	1	US-06-129-704-2
40	59	9.3	114	1	US-06-129-704-2
41	59	9.3	114	1	US-06-129-704-2
42	59	9.3	114	1	US-06-129-704-2
43	59	9.3	114	1	US-06-129-704-2
44	59	9.3	114	1	US-06-129-704-2
45	59	9.3	114	1	US-06-129-704-2

## ALIGNMENTS

RESULT 1

Sequence 2, Application PC/US9507171

GENERAL INFORMATION:

APPLICANT: MCJNEY, J.

TITLE OF INVENTION: Human Chemokine Beta-12

NUMBER OF SEQUENCES: 7

CORRESPONDENT ADDRESS:

ADDRESSEE: CARILIA, BYRN, BAIN, GILFILLAN,

SIGNER: BOCHI, STEWART & OLSTEIN

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PC/US95/07171

FILING DATE: 6 JUN 95

CLASSIFICATION:

Sequence 8, Appl	Sequence 9, Appl	Sequence 10, Appl	Sequence 11, Appl	Sequence 12, Appl	Sequence 13, Appl	Sequence 14, Appl	Sequence 15, Appl	Sequence 16, Appl	Sequence 17, Appl	Sequence 18, Appl	Sequence 19, Appl	Sequence 20, Appl	Sequence 21, Appl	Sequence 22, Appl	Sequence 23, Appl	Sequence 24, Appl	Sequence 25, Appl	Sequence 26, Appl	Sequence 27, Appl	Sequence 28, Appl	Sequence 29, Appl	Sequence 30, Appl	Sequence 31, Appl	Sequence 32, Appl	Sequence 33, Appl	Sequence 34, Appl	Sequence 35, Appl	Sequence 36, Appl	Sequence 37, Appl	Sequence 38, Appl	Sequence 39, Appl	Sequence 40, Appl	Sequence 41, Appl	Sequence 42, Appl	Sequence 43, Appl	Sequence 44, Appl	Sequence 45, Appl	Sequence 46, Appl	Sequence 47, Appl	Sequence 48, Appl	Sequence 49, Appl	Sequence 50, Appl	Sequence 51, Appl	Sequence 52, Appl	Sequence 53, Appl	Sequence 54, Appl	Sequence 55, Appl	Sequence 56, Appl	Sequence 57, Appl	Sequence 58, Appl	Sequence 59, Appl	Sequence 60, Appl	Sequence 61, Appl	Sequence 62, Appl	Sequence 63, Appl	Sequence 64, Appl	Sequence 65, Appl	Sequence 66, Appl	Sequence 67, Appl	Sequence 68, Appl	Sequence 69, Appl	Sequence 70, Appl	Sequence 71, Appl	Sequence 72, Appl	Sequence 73, Appl	Sequence 74, Appl	Sequence 75, Appl	Sequence 76, Appl	Sequence 77, Appl	Sequence 78, Appl	Sequence 79, Appl	Sequence 80, Appl	Sequence 81, Appl	Sequence 82, Appl	Sequence 83, Appl	Sequence 84, Appl	Sequence 85, Appl	Sequence 86, Appl	Sequence 87, Appl	Sequence 88, Appl	Sequence 89, Appl	Sequence 90, Appl	Sequence 91, Appl	Sequence 92, Appl	Sequence 93, Appl	Sequence 94, Appl	Sequence 95, Appl	Sequence 96, Appl	Sequence 97, Appl	Sequence 98, Appl	Sequence 99, Appl	Sequence 100, Appl
------------------	------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	--------------------

LENGTH: 136 AMINO ACIDS  
 TYPE: AMINO ACID  
 STRANDEDNESS:  
 TOPOLOGY: LINEAR  
 MOLECULE TYPE: PROTEIN  
 PCT-US95-07171-2

## Alignment Scores:

Pred. No.:	0.00209	Length:	136
Score:	83.00	Matches:	24
Percent Similarity:	42.86%	Conservative:	15
Best Local Similarity:	26.37%	Mismatches:	32
Query Match:	12.63%	Indels:	20
DB:	5	Gaps:	3

US-09-762-027-4 (1-388) x PCT-US95-07171-2 (1-136)

```

QY 136 CCGAGCTCGCTGATTTATTTGTTTCAGTCCTAGAGATATAGTAAAGAACTT 135
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5 AyleuSerLeuLysLeuIrycysLeuAlaGluAlaGlyMetLysValSerGluAlaSer 24
QY 196 TCCATTCGCCTACTG----- 210
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25 SerLeuLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 44
QY 211 -----AGAACTGCGAGCGAGCGAGCTGCTAGAGCTTAC---AAAGAACTTAAAT 261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 45 GluTrpValAsnThrProSerThrCysLeuLysTyrTyrGluLysValLeuProArg 64
QY 262 CATCTTCTTGAGCGA-----AAGAACTGAACTGCTACATAGAGAGCTTGCTTCAATA 315
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 ArgLeuValValLysIryArgLysAlaLeuAsnCysHisLeuProAlaLeuLeuPheVal 64
QY 316 ACTCTCAGAACTGCGAGATTCCTGTATAATCCT 348
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 ThrLysArgAsnArgGluValLysThrAsnPro 95
  
```

## RESULT 2

PCT-US95-07171-3  
 Sequence 3 Application PCT/US9507171

GENERAL INFORMATION:

APPLICANT: MOONEY, J.

TITLE OF INVENTION: Human Chemokine Beta-12

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESS: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/07171

FILING DATE: 6 JUN 95

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: None

FILING DATE: None

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-391

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 136 AMINO ACIDS  
 TYPE: AMINO ACID  
 STRANDEDNESS:  
 TOPOLOGY: LINEAR  
 MOLECULE TYPE: PROTEIN  
 PCT-US95-07171-2

## Alignment Scores:

Pred. No.:	0.00209	Length:	136
Score:	83.00	Matches:	24
Percent Similarity:	42.86%	Conservative:	15
Best Local Similarity:	26.37%	Mismatches:	32
Query Match:	12.63%	Indels:	20
DB:	5	Gaps:	3

US-09-762-027-4 (1-388) x PCT-US95-07171-3 (1-136)

```

QY 136 CCGAGCTCGCTGATTTATTTGTTTCAGTCCTAGAGATATAGTAAAGAACTT 135
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5 AyleuSerLeuLysLeuIrycysLeuAlaGluAlaGlyMetLysValSerGluAlaSer 24
QY 196 TCCATTCGCCTACTG----- 210
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25 SerLeuLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 44
QY 211 -----AGAACTGCGAGCGAGCGAGCTGCTAGAGCTTAC---AAAGAACTTAAAT 261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 45 GluTrpValAsnThrProSerThrCysLeuLysTyrTyrGluLysValLeuProArg 64
QY 262 CATCTTCTTGAGCGA-----AAGAACTGAACTGCTACATAGAGAGCTTGCTTCAATA 315
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 ArgLeuValValLysIryArgLysAlaLeuAsnCysHisLeuProAlaLeuLeuPheVal 64
QY 316 ACTCTCAGAACTGCGAGATTCCTGTATAATCCT 348
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 ThrLysArgAsnArgGluValLysThrAsnPro 95
  
```

## RESULT 3

US-07-977-434  
 Sequence 3 Application US/07977434

GENERAL INFORMATION:

APPLICANT: Gelland, Barry H.

TITLE OF INVENTION: 5' to 3' LYMPHOCYTE MUTATIONS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Gelland, Barry H.

STREET: 310 Kingsland Street

CITY: New Jersey

STATE: New Jersey

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07977434

FILING DATE: 05/07/97

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 590,190

FILING DATE: 20-SEP-1990

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-391

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:





```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

US-06-222-017A
Alignment S: 1965
Field: 101
Score: 44
Percent Similarity: 37.6
Best Local Similarity: 24.7
Query Matrix: 2
Gap(s): 5

US-06-222-017A (1-3665)
1 355 2 355 3 355 4 355 5 355 6 355 7 355 8 355 9 355 10 355 11 355 12 355 13 355 14 355 15 355 16 355 17 355 18 355 19 355 20 355 21 355 22 355 23 355 24 355 25 355 26 355 27 355 28 355 29 355 30 355 31 355 32 355 33 355 34 355 35 355 36 355 37 355 38 355 39 355 40 355 41 355 42 355 43 355 44 355 45 355 46 355 47 355 48 355 49 355 50 355 51 355 52 355 53 355 54 355 55 355 56 355 57 355 58 355 59 355 60 355 61 355 62 355 63 355 64 355 65 355 66 355 67 355 68 355 69 355 70 355 71 355 72 355 73 355 74 355 75 355 76 355 77 355 78 355 79 355 80 355 81 355 82 355 83 355 84 355 85 355 86 355 87 355 88 355 89 355 90 355 91 355 92 355 93 355 94 355 95 355 96 355 97 355 98 355 99 355 100 355 101 355 102 355 103 355 104 355 105 355 106 355 107 355 108 355 109 355 110 355 111 355 112 355 113 355 114 355 115 355 116 355 117 355 118 355 119 355 120 355 121 355 122 355 123 355 124 355 125 355 126 355 127 355 128 355 129 355 130 355 131 355 132 355 133 355 134 355 135 355 136 355 137 355 138 355 139 355 140 355 141 355 142 355 143 355 144 355 145 355 146 355 147 355 148 355 149 355 150 355 151 355 152 355 153 355 154 355 155 355 156 355 157 355 158 355 159 355 160 355 161 355 162 355 163 355 164 355 165 355 166 355 167 355 168 355 169 355 170 355 171 355 172 355 173 355 174 355 175 355 176 355 177 355 178 355 179 355 180 355 181 355 182 355 183 355 184 355 185 355 186 355 187 355 188 355 189 355 190 355 191 355 192 355 193 355 194 355 195 355 196 355 197 355 198 355 199 355 200 355 201 355 202 355 203 355 204 355 205 355 206 355 207 355 208 355 209 355 210 355 211 355 212 355 213 355 214 355 215 355 216 355 217 355 218 355 219 355 220 355 221 355 222 355 223 355 224 355 225 355 226 355 227 355 228 355 229 355 230 355 231 355 232 355 233 355 234 355 235 355 236 355 237 355 238 355 239 355 240 355 241 355 242 355 243 355 244 355 245 355 246 355 247 355 248 355 249 355 250 355 251 355 252 355 253 355 254 355 255 355 256 355 257 355 258 355 259 355 260 355 261 355 262 355 263 355 264 355 265 355 266 355 267 355 268 355 269 355 270 355 271 355 272 355 273 355 274 355 275 355 276 355 277 355 278 355 279 355 280 355 281 355 282 355 283 355 284 355 285 355 286 355 287 355 288 355 289 355 290 355 291 355 292 355 293 355 294 355 295 355 296 355 297 355 298 355 299 355 300 355 301 355 302 355 303 355 304 355 305 355 306 355 307 355 308 355 309 355 310 355 311 355 312 355 313 355 314 355 315 355 316 355 317 355 318 355 319 355 320 355 321 355 322 355 323 355 324 355 325 355 326 355 327 355 328 355 329 355 330 355 331 355 332 355 333 355 334 355 335 355 336 355 337 355 338 355 339 355 340 355 341 355 342 355 343 355 344 355 345 355 346 355 347 355 348 355 349 355 350 355 351 355 352 355 353 355 354 355 355 356 355 357 355 358 355 359 355 360 355 361 355 362 355 363 355 364 355 365 355 366 355 367 355 368 355 369 355 370 355 371 355 372 355 373 355 374 355 375 355 376 355 377 355 378 355 379 355 380 355 381 355 382 355 383 355 384 355 385 355 386 355 387 355 388 355 389 355 390 355 391 355 392 355 393 355 394 355 395 355 396 355 397 355 398 355 399 355 400 355 401 355 402 355 403 355 404 355 405 355 406 355 407 355 408 355 409 355 410 355 411 355 412 355 413 355 414 355 415 355 416 355 417 355 418 355 419 355 420 355 421 355 422 355 423 355 424 355 425 355 426 355 427 355 428 355 429 355 430 355 431 355 432 355 433 355 434 355 435 355 436 355 437 355 438 355 439 355 440 355 441 355 442 355 443 355 444 355 445 355 446 355 447 355 448 355 449 355 450 355 451 355 452 355 453 355 454 355 455 355 456 355 457 355 458 355 459 355 460 355 461 355 462 355 463 355 464 355 465 355 466 355 467 355 468 355 469 355 470 355 471 355 472 355 473 355 474 355 475 355 476 355 477 355 478 355 479 355 480 355 481 355 482 355 483 355 484 355 485 355 486 355 487 355 488 355 489 355 490 355 491 355 492 355 493 355 494 355 495 355 496 355 497 355 498 355 499 355 500 355 501 355 502 355 503 355 504 355 505 355 506 355 507 355 508 355 509 355 510 355 511 355 512 355 513 355 514 355 51
```



TITLE OF INVENTION: EXPRESSED CHEMOKINES, THEIR  
TITLE OF INVENTION: PRODUCTION AND USES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSP Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/347,492B  
FILING DATE: 29-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/303,241  
FILING DATE: 07-SEP-1994  
APPLICATION NUMBER: 08/320,011  
FILING DATE: 05-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Luther, Barbara J  
REGISTRATION NUMBER: 33,954  
REFERENCE/DOCKET NUMBER: PF-0024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-852-0195  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: LIVER  
CLONE: 87825  
US-09-762-027-4 (1-388) x US-08-347-492B-2 (1-120)  
Alignment Scores:  
Pred. No.: 1 23 Length: 120  
Score: 63.50 Matches: 16  
Percent Similarity: 47.92% Conservative: 7  
Best Local Similarity: 33.33% Mismatches: 22  
Query Match: 9.67% Indels: 3  
Gaps: 2  
US-09-762-027-4 (1-388) x US-08-347-492B-2 (1-120)  
214 AACTGGCAGCGAGGTGCTACAGGCTAC---AAAGAACTTCAATCATCTCTT 27  
111 111 111 111  
32 AsnThrProSerThyCysCysLeuLysIleValLeuProCysArgSerVal 51  
111 111 111 111 111 111 111 111  
271 GAGGGA-----AAGAACTGAGTGTACATCAATGCTGCTGTTATACCTTCA 344  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
52 ValGlyTyrArgLysAlaLeuAsnGlyHisLeuProAlaIleIlePheValIleLysArg 71  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
325 AGCTGCAGATTCTGTATTAATCCT 348  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
72 AsnArgGluValCysThrAsnPro 79  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
RESULT 10  
US-08-798-143-2  
Sequence 2, Application US/08798143  
Patent No. 5936068  
GENERAL INFORMATION:  
APPLICANT: Wilde, Craig G.  
APPLICANT: Hawkins, Phillip R.  
APPLICANT: Bandman, Olga  
APPLICANT: Sellhammer, Jeffrey J.

TITLE OF INVENTION: EXPRESSED CHEMOKINES, THEIR  
TITLE OF INVENTION: PRODUCTION AND USES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSP Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/798,143  
FILING DATE: 11-FEB-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/347,492  
FILING DATE: 29-NOV-1994  
APPLICATION NUMBER: 08/303,241  
FILING DATE: 07-SEP-1994  
APPLICATION NUMBER: 08/320,011  
FILING DATE: 05-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Luther, Barbara J  
REGISTRATION NUMBER: 33,954  
REFERENCE/DOCKET NUMBER: PF-0024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-852-0195  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: LIVER  
CLONE: 87825  
US-09-762-027-4 (1-388) x US-08-798-143-2 (1-120)  
Alignment Scores:  
Pred. No.: 1 23 Length: 120  
Score: 63.50 Matches: 16  
Percent Similarity: 47.92% Conservative: 7  
Best Local Similarity: 33.33% Mismatches: 22  
Query Match: 9.67% Indels: 3  
Gaps: 2  
US-09-762-027-4 (1-388) x US-08-798-143-2 (1-120)  
214 AACTGGCAGCGAGGTGCTACAGGCTAC---AAAGAACTTCAATCATCTCTT 270  
111 111 111 111 111 111 111 111  
32 AsnThrProSerThyCysCysLeuLysIleValLeuProCysArgSerVal 51  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
271 GAGGGA-----AAGAACTGAGTGTACATCAATGCTGCTGTTATACCTTCA 344  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
52 ValGlyTyrArgLysAlaLeuAsnGlyHisLeuProAlaIleIlePheValIleLysArg 71  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
325 AGCTGCAGATTCTGTATTAATCCT 348  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
72 AsnArgGluValCysThrAsnPro 79  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
RESULT 11  
US-09-762-027-4 (1-388) x US-08-798-143-2 (1-120)  
Sequence 2, Application US/08798143  
Patent No. 5936068  
GENERAL INFORMATION:  
APPLICANT: Wilde, Craig G.  
APPLICANT: Hawkins, Phillip R.  
APPLICANT: Bandman, Olga  
APPLICANT: Sellhammer, Jeffrey J.









GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

us nucleic - protein search, using frame\_plus\_n2p model

Run on: January 2, 2003, 15:08:03 : Search time 49 seconds  
(without alignment)

735,319 Million cell updates/sec

File: US-09-762-027-4  
Perfect score: 657

Sequence: 1 agctgcatacaagaaacac.....atccataaagtatattatccatcagc

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Zgapop 6.0	Zgapext 7.0
Delop 6.0	Delext 7.0

Searched: 117078 seqs, 18951520 residues

Local number of hits satisfying chosen parameters: 41156

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_n2p\_model -DEV=xlh  
-Q=/cgn2\_1/uspro.spool/US09762027/runal\_02012003\_085614\_2632/epc\_qntr7.fasta.1.53  
-DB=published.Applications\_AA -QEMF=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-IRANS=human0.cdi -DIST=45 -DOCCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NCOR=ext -HEAPSIZ=500 -MINLEN=2000000000 -USER=US09762027\_scon.1.1.2 -runat\_02012003\_085614\_2632  
-NCPU=6 -ICPU=3 -NO\_MLPHY -NO\_MMAPP -LARGEOUTPUT -NEE\_SCORES=0 -MATT=1 -LARGEST3  
-DEV\_1TIMEOUT=120 -WARN\_ETIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=7  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications AA:

- 1: /cgn2\_6/prodata/2/pubppa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/prodata/2/pubppa/PC1\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/2/pubppa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/2/pubppa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/prodata/2/pubppa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/2/pubppa/PCUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/prodata/2/pubppa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/2/pubppa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/2/pubppa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/prodata/2/pubppa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/prodata/2/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	13.9	241	10	US-09-925-300-1065 Sequence 1065, App1
2	77	11.7	181	10	US-09-867-950-930 Sequence 930, App1
3	61.5	9.4	372	10	US-09-768-894A-2 Sequence 2, App1
4	60	9.1	276	10	US-09-943-798-2 Sequence 2, App1

1	91	13.9	241	10	US-09-943-798-4 Sequence 4, App1
2	77	11.7	181	10	US-09-864-761-18727 Sequence 18727, A
3	61.5	9.4	372	10	US-09-771-956-20 Sequence 20, App1
4	60	9.1	276	10	US-09-771-956-6 Sequence 6, App1
5	59	9.0	352	10	US-09-771-956-23 Sequence 23, App1
6	58	8.9	348	10	US-09-745-21 Sequence 41, App1
7	57	8.8	341	10	US-09-771-956-10 Sequence 10, App1
8	56	8.7	339	10	US-09-771-956-25 Sequence 25, App1
9	55	8.6	325	10	US-09-870-574-3 Sequence 3, App1
10	54	8.5	325	10	US-09-949-192-5 Sequence 5, App1
11	53	8.4	325	10	US-10-062-586-350 Sequence 350, App
12	52	8.3	325	10	US-09-815-242-10574 Sequence 10574, A
13	51	8.2	325	10	US-09-815-242-10574 Sequence 10574, A
14	50	8.1	325	10	US-09-815-242-10574 Sequence 10574, A
15	49	8.0	325	10	US-09-815-242-10574 Sequence 10574, A
16	48	7.9	325	10	US-09-815-242-10574 Sequence 10574, A
17	47	7.8	325	10	US-09-815-242-10574 Sequence 10574, A
18	46	7.7	325	10	US-09-815-242-10574 Sequence 10574, A
19	45	7.6	325	10	US-09-815-242-10574 Sequence 10574, A
20	44	7.5	325	10	US-09-815-242-10574 Sequence 10574, A
21	43	7.4	325	10	US-09-815-242-10574 Sequence 10574, A
22	42	7.3	325	10	US-09-815-242-10574 Sequence 10574, A
23	41	7.2	325	10	US-09-815-242-10574 Sequence 10574, A
24	40	7.1	325	10	US-09-815-242-10574 Sequence 10574, A
25	39	7.0	325	10	US-09-815-242-10574 Sequence 10574, A
26	38	6.9	325	10	US-09-815-242-10574 Sequence 10574, A
27	37	6.8	325	10	US-09-815-242-10574 Sequence 10574, A
28	36	6.7	325	10	US-09-815-242-10574 Sequence 10574, A
29	35	6.6	325	10	US-09-815-242-10574 Sequence 10574, A
30	34	6.5	325	10	US-09-815-242-10574 Sequence 10574, A
31	33	6.4	325	10	US-09-815-242-10574 Sequence 10574, A
32	32	6.3	325	10	US-09-815-242-10574 Sequence 10574, A
33	31	6.2	325	10	US-09-815-242-10574 Sequence 10574, A
34	30	6.1	325	10	US-09-815-242-10574 Sequence 10574, A
35	29	6.0	325	10	US-09-815-242-10574 Sequence 10574, A
36	28	5.9	325	10	US-09-815-242-10574 Sequence 10574, A
37	27	5.8	325	10	US-09-815-242-10574 Sequence 10574, A
38	26	5.7	325	10	US-09-815-242-10574 Sequence 10574, A
39	25	5.6	325	10	US-09-815-242-10574 Sequence 10574, A
40	24	5.5	325	10	US-09-815-242-10574 Sequence 10574, A
41	23	5.4	325	10	US-09-815-242-10574 Sequence 10574, A
42	22	5.3	325	10	US-09-815-242-10574 Sequence 10574, A
43	21	5.2	325	10	US-09-815-242-10574 Sequence 10574, A
44	20	5.1	325	10	US-09-815-242-10574 Sequence 10574, A
45	19	5.0	325	10	US-09-815-242-10574 Sequence 10574, A

## ALIGNMENTS

Result 1  
US-09-925-300-1065  
Sequence 1065, App1  
General Info: Application: US09-925-300  
Applicant: Steve Rosen  
Title of Invention: Nucleic Acids, Proteins and Antibodies  
Current Application No: US09-925-300  
Current Filing Date: 1999-03-10  
Prior Filing Date: 1999-03-10  
Prior Application Number: 60/124,270  
Priority Filing Date: 1999-03-12  
Number of Seq ID Nos: 1690  
Strainer: Pat. Status Ver: 2.0  
Seq ID No: 1065  
Length: 241  
Type: P1  
Organism: Homo sapiens  
US-09-925-300-1065  
Alignment Scores:  
Pred. No.: 91  
Score: 13.9  
Percent Similarity: 91.00  
Best Local Similarity: 91.14  
Length: 241  
Matches: 20  
Conservative: 7  
Mismatches: 7

```

Query Match: 13.65%   Models: 1
DB: 10               Gaps: 0
US-09-762-027-4 (1-388) x US-09-925-300-1065 (-211)

QY 282 GCTGAAGTCTCATAGATGCTGTGCTCAATAAATCAGAGCTGAGATTCATAT 341
DB 9 SerGilaSerHisArgMetIleLeuLeuAsnSerHisLysLeuLeuAlaLeuTyr 26
QY 342 AAATCCTGAAAACAGACATCCCTGATCATCAATAAGTATAT 384
DB 29 LysSerLeuAlaArgSerIle-ProLysLeuLysValTyr 42

RESULT 2
US-09-867-550-930
Sequence 950, Application US/09867550
Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Foad.
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atheroarterial cells and
FILE REFERENCE: 21402-013 (Coura-313)
CURRENT APPLICATION NUMBER: US/09867,550
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: USSN 60/208,427
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 930
LENGTH: 181
TYPE: PRI
ORGANISM: Homo sapiens
US-09-867-550-930

Alignment Scores:
Pred. No.: 0.0156   Length: 181
Score: 77.60       Matches: 16
Percent Similarity: 79.31%   Conservative: 5
Best Local Similarity: 62.07%   Mismatches: 5
Query Match: 11.72%   Indels: 1
DB: 10             Gaps: 0

US-09-762-027-4 (1-388) x US-09-867-550-930 (1-181)

QY 300 ATGCTTGCTGCTCATACACTGACAGAGCTGCAGATTCATATAATCTTACGAAATAT 359
DB 1 MetIleLeuLeuAsnSerHisLysLysLeuAlaLeuTyrLysSerIleAlaAlaTser 2
QY 360 ATCCCTGATCATCAATAAGTATAT 384
DB 21 Ile-ProLysLeuLysValTyr 28

RESULT 3
US-09-768-894A-2
Sequence 2, Application US/09768894A
Patent No. US20020115102A1
GENERAL INFORMATION:
APPLICANT: Diane Joan Couzens
APPLICANT: Steven Michael Foord
APPLICANT: Diane Michele Ignar
APPLICANT: Filippo Volpe
TITLE OF INVENTION: NOVEL PROTEIN
FILE REFERENCE: PG377005
CURRENT APPLICATION NUMBER: US/09768,894A
CURRENT FILING DATE: 2002-02-24
PRIOR APPLICATION NUMBER: 09/768,894
PRIOR FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1

```

```

SEQ ID NO 930
LENGTH: 181
TYPE: PRI
ORGANISM: Homo sapiens
US-09-768-894A-2

Alignment Scores:
Pred. No.: 0.0156   Length: 181
Score: 77.60       Matches: 16
Percent Similarity: 79.31%   Conservative: 5
Best Local Similarity: 62.07%   Mismatches: 5
Query Match: 11.72%   Indels: 1
DB: 10             Gaps: 0

US-09-762-027-4 (1-388) x US-09-867-550-930 (1-181)

QY 300 ATGCTTGCTGCTCATACACTGACAGAGCTGCAGATTCATATAATCTTACGAAATAT 359
DB 1 MetIleLeuLeuAsnSerHisLysLysLeuAlaLeuTyrLysSerIleAlaAlaTser 2
QY 360 ATCCCTGATCATCAATAAGTATAT 384
DB 21 Ile-ProLysLeuLysValTyr 28

RESULT 3
US-09-768-894A-2
Sequence 2, Application US/09768894A
Patent No. US20020115102A1
GENERAL INFORMATION:
APPLICANT: Diane Joan Couzens
APPLICANT: Steven Michael Foord
APPLICANT: Diane Michele Ignar
APPLICANT: Filippo Volpe
TITLE OF INVENTION: NOVEL PROTEIN
FILE REFERENCE: PG377005
CURRENT APPLICATION NUMBER: US/09768,894A
CURRENT FILING DATE: 2002-02-24
PRIOR APPLICATION NUMBER: 09/768,894
PRIOR FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1

```

Db 259 LysValSerGlyAsnLeuGluGlnAlaLysLysIle 270

RESULT 5

US-09-943-798-4  
Sequence 1, Application US/09943798  
Patent No. US20020065215A1  
GENERAL INFORMATION:  
APPLICANT: Glaxo Group Limited  
TITLE OF INVENTION: Polypeptide  
FILE REFERENCE: Q61021  
CURRENT APPLICATION NUMBER: US/09/943,798  
CURRENT FILING DATE: 2001-08-31  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 337  
TYPE: PRI  
ORGANISM: Homo sapiens  
US-09-943-798-4

# Alignment Scores:

Seq. No.:	Length:	Matches:	Score:
1	337	20	60.00
2	20	7	51.928
3	17	8	38.468
4	8	1	5.138
5	2	1	10

US-09-762-027-4 (1-388) x US-09-943-798-4 (1-337)

QY 148 GAATTGATGCTTTCAGTCCGACAGATATTCACACAGAAC-----192  
Db 282 GlnAlaIyrIleValSer-----ArgProLeuAlaIleValSerThrPhcGlyAsnLeu 299  
QY 193 ---CTTTCATGCTGCTACTGAGAACTGGCAGCAGGAGCTGCTCCTACAGCTACAAA 243  
Db 300 LeuLeuIyrValIleValSerAspAsnPhcGlnAlaValLysSerThrValArgGlys 319  
QY 250 GAACCTGACATCTCTCTGAGGAGGAGACGCT 285  
Db 320 LysValSerGlyAsnLeuGluGlnAlaLysLysIle 331

# RESULT 6

US-09-864-761-48727  
Sequence 18727, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROCESS USEFUL FOR  
FILE REFERENCE: A600000000-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30

US-09-864-761-48727 x US-09-943-798-4 (1-337)  
QY 56 GATGATGCTTTCAGTCCGACAGATATTCACACAGAAC-----192  
Db 282 GlnAlaIyrIleValSer-----ArgProLeuAlaIleValSerThrPhcGlyAsnLeu 299  
QY 193 ---CTTTCATGCTGCTACTGAGAACTGGCAGCAGGAGCTGCTCCTACAGCTACAAA 243  
Db 300 LeuLeuIyrValIleValSerAspAsnPhcGlnAlaValLysSerThrValArgGlys 319  
QY 250 GAACCTGACATCTCTCTGAGGAGGAGACGCT 285  
Db 320 LysValSerGlyAsnLeuGluGlnAlaLysLysIle 331

# Alignment Scores:

Seq. No.:	Length:	Matches:	Score:
1	337	20	60.00
2	20	7	51.928
3	17	8	38.468
4	8	1	5.138
5	2	1	10

US-09-864-761-48727 x US-09-943-798-4 (1-337)

QY 159 -----CGTACAGTATAGCTACAGAACCTTCC-----196  
Db 120 GlnAlaIyrIleValSerAspAsnPhcGlnAlaValLysSerThrValArgGlys 319  
QY 193 ---CTTTCATGCTGCTACTGAGAACTGGCAGCAGGAGCTGCTCCTACAGCTACAAA 243  
Db 300 LeuLeuIyrValIleValSerAspAsnPhcGlnAlaValLysSerThrValArgGlys 319  
QY 250 GAACCTGACATCTCTCTGAGGAGGAGACGCT 285  
Db 320 LysValSerGlyAsnLeuGluGlnAlaLysLysIle 331

# RESULT 7

US-09-771-95574  
Sequence 20, Appl. No. US/09771956  
Patent No. US 6,147,761  
GENERAL INFORMATION:  
APPLICANT: Knudsen, James  
APPLICANT: Knudsen, James  
TITLE OF INVENTION: Human Neuropeptide Y Receptors  
FILE REFERENCE: A600000000-1  
CURRENT APPLICATION NUMBER: US/09/771,956



```

Db      268  tPolehniSvalPhehniSvalValaThrAsprhEasnsAspnsleuLleSerFasnaatJhl 1
OY      262  GCtAAAGTCACACTAGAGATGCTGTGCC---TTCAATACCTCTCAAGAGTTCACATATT 262
Db      308  sPhelys---LeuValHyPCysHicCysHISleuLeuGlyMetMetSerYs-----Gy 308
OY      339  TATAATCTCT 348
Db      325  sleuAsnpro 328

RESULT 10
US-09-909-745-23
; Sequence 23, Application US/09909745
; Patent NO. US20020119546A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; TITLE OF INVENTION: Squalene Synthesis Enzymes
; FILE REFERENCE: BB1112 US CIP
; CURRENT APPLICATION NUMBER: US/09/909,745
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-909-745-23

Alignment Scores:
Pred. No.:      7.47      length:      376
Score:          59.00      Matches:      20
Percent Similarity: 41.10%      Conservative: 10
Best Local Similarity: 27.40%      Mismatches: 25
Query Match:      8.99%      Indels:      16
DB:              10      gaps:      2

US-09-752-027-4 (1-388) x US-09-909-745-23 (1-378)
OY      151  CTATACTCTACGCGACTGAAACAAATTAATTCAGTACGCTGTTGACTGTTAG 151
Db      305  LeuLeuSerMetGlyValSerHisSerHisSerHisleuValaIleleuThrThrVallys 305
OY      121  CATAA-----GCTTCAGCAGATTATGGTTCGTCGACATTTAA 121
Db      245  HISLysleuValSerLysleuThrGlyAlaGlyIleGlyGlyValIleuIleuIleu 245
OY      82  CCAGACGTTCTGTCATAGACAGAAATACACAGCGCTTTGTAGACAGAA 82
Db      345  ProThrGlyThrValValaAspLysValValaGluGluIleuGluIleuSerSerGlyPheHisGly 345
OY      31  -----GACTGCAGATATGT 17
Db      365  PheThrAlaLeuIleGlyGlyAsnGlyAlaGluIleIleCys 377

RESULT 11
US-09-771-956-21
; Sequence 21, Application US/09771956
; Patent NO. US20010031474A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Michale
; APPLICANT: Brodbeck, Robin
; APPLICANT: Krause, James
; TITLE OF INVENTION: Chimeric Neuropeptide Y Receptors
; FILE REFERENCE: N2000.001
; CURRENT APPLICATION NUMBER: US/09/771,956
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 31

```

```

US-07-771-951
Seq. ID No. 1
Length: 383
Type: PE
ORIGIN: Official Sequence
FEATURE:
  OTHER IN: MATCH: Description of Artificial Sequence: Y/Y1 CHIMERA
US-07-771-951
Alignment Score:
  Pred. No.: 7.5
  Score: 59.00
  Percent Similarity: 41.35%
  Best Local Similarity: 25.00%
  Query Match: 6.96%
  Ds: 10
  Caps: 7
US-07-771-951 (1-368) x US-07-771-956-21 (1-383)
Q1 117 TAACTTAACTAGTACACTTAACTAGTAAATTTATGCTTTCAAGCCAGAGT 176
   117 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
D1 225 VET TLEUHTVALSERHISIMSERVAL-----CYSLEARGLEUVALSAG-- 240
Q2 177 AATAAACAACAAC----- 152
   117 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
D2 241 -----ALGSHASAMTETASPLYSMETARGASPSNLSYIATFASERAT 257
Q3 153 -----TTTTCATTCGCACACTG-----AAGACCTGGCA 221
   153 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
D3 257 GATTTGATSERVALPHETATLDELHETLEUHTLEUVALPHEALVALSERITPHE 277
Q4 242 GATTTTACTTGCTTACCTGTTTAAATTAAGAACTTCAGATGATCTTATGCTTAAAGAA 261
   242 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
D4 277 TPT TLEUHTSLEUPHEHSVALTHTASPSHASPASPSNLTLESTASATAGHI 297
Q5 252 GCTTAAGTACATAAAGCTTGCTGCTCTTCTACTCTCGAAGACGTCCAGATCTG 338
   252 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
D5 257 SPNLS-----TENVATLYTGYSTLECYSHISLEUHTLEUHTMETCETSCYS-----CY 314
Q6 359 TATCATCTC 348
   359 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
D6 314 STGASPTPC 317
RESULT 12
US-07-771-951
  Sequence: Application US-0771952
  Patent No.: 526010014741
  General Information:
    APPLICANT: Bennett, Michael
    ATTORNEY: Brobeck, Robert
    FILE NO. APPLICATION: Chimeric Neuropeptide Y Receptors
    FILE REFERENCE: R2000.004
    CURRENT APPLICATION NUMBER: US-0771-955
    CURRENT FILING DATE: 2001-1-25
    NUMBER OF SEQ. ID NOS.: 31
    S. NUMBER: Patent Ver. 2.1
    SEQ. ID NO.:
      1:
      2:
      3:
      4:
      5:
      6:
      7:
      8:
      9:
      10:
      11:
      12:
      13:
      14:
      15:
      16:
      17:
      18:
      19:
      20:
      21:
      22:
      23:
      24:
      25:
      26:
      27:
      28:
      29:
      30:
      31:
      32:
      33:
      34:
      35:
      36:
      37:
      38:
      39:
      40:
      41:
      42:
      43:
      44:
      45:
      46:
      47:
      48:
      49:
      50:
      51:
      52:
      53:
      54:
      55:
      56:
      57:
      58:
      59:
      60:
      61:
      62:
      63:
      64:
      65:
      66:
      67:
      68:
      69:
      70:
      71:
      72:
      73:
      74:
      75:
      76:
      77:
      78:
      79:
      80:
      81:
      82:
      83:
      84:
      85:
      86:
      87:
      88:
      89:
      90:
      91:
      92:
      93:
      94:
      95:
      96:
      97:
      98:
      99:
      100:
      101:
      102:
      103:
      104:
      105:
      106:
      107:
      108:
      109:
      110:
      111:
      112:
      113:
      114:
      115:
      116:
      117:
      118:
      119:
      120:
      121:
      122:
      123:
      124:
      125:
      126:
      127:
      128:
      129:
      130:
      131:
      132:
      133:
      134:
      135:
      136:
      137:
      138:
      139:
      140:
      141:
      142:
      143:
      144:
      145:
      146:
      147:
      148:
      149:
      150:
      151:
      152:
      153:
      154:
      155:
      156:
      157:
      158:
      159:
      160:
      161:
      162:
      163:
      164:
      165:
      166:
      167:
      168:
      169:
      170:
      171:
      172:
      173:
      174:
      175:
      176:
      177:
      178:
      179:
      180:
      181:
      182:
      183:
      184:
      185:
      186:
      187:
      188:
      189:
      190:
      191:
      192:
      193:
      194:
      195:
      196:
      197:
      198:
      199:
      200:
      201:
      202:
      203:
      204:
      205:
      206:
      207:
      208:
      209:
      210:
      211:
      212:
      213:
      214:
      215:
      216:
      217:
      218:
      219:
      220:
      221:
      222:
      223:
      224:
      225:
      226:
      227:
      228:
      229:
      230:
      231:
      232:
      233:
      234:
      235:
      236:
      237:
      238:
      239:
      240:
      241:
      242:
      243:
      244:
      245:
      246:
      247:
      248:
      249:
      250:
      251:
      252:
      253:
      254:
      255:
      256:
      257:
      258:
      259:
      260:
      261:
      262:
      263:
      264:
      265:
      266:
      267:
      268:
      269:
      270:
      271:
      272:
      273:
      274:
      275:
      276:
      277:
      278:
      279:
      280:
      281:
      282:
      283:
      284:
      285:
      286:
      287:
      288:
      289:
      290:
      291:
      292:
      293:
      294:
      295:
      296:
      297:
      298:
      299:
      300:
      301:
      302:
      303:
      304:
      305:
      306:
      307:
      308:
      309:
      310:
      311:
      312:
      313:
      314:
      315:
      316:
      317:
      318:
      319:
      320:
      321:
      322:
      323:
      324:
      325:
      326:
      327:
      328:
      329:
      330:
      331:
      332:
      333:
      334:
      335:
      336:
      337:
      338:
      339:
      340:
      341:
      342:
      343:
      344:
      345:
      346:
      347:
      348:
      349:
      350:
      351:
      352:
      353:
      354:
      355:
      356:
      357:
      358:
      359:
      360:
      361:
      362:
      363:
      364:
      365:
      366:
      367:
      368:
      369:
      370:
      371:
      372:
      373:
      374:
      375:
      376:
      377:
      378:
      379:
      380:
      381:
      382:
      383:
      384:
      385:
      386:
      387:
      388:
      389:
      390:
      391:
      392:
      393:
      394:
      395:
      396:
      397:
      398:
      399:
      400:
      401:
      402:
      403:
      404:
      405:
      406:
      407:
      408:
      409:
      410:
      411:
      412:
      413:
      414:
      415:
      416:
      417:
      418:
      419:
      420:
      421:
      422:
      423:
      424:
      425:
      426:
      427:
      428:
      429:
      430:
      431:
      432:
      433:
      434:
      435:
      436:
      437:
      438:
      439:
      440:
      441:
      442:
      443:
      444:
      445:
      446:
      447:
      448:
      449:
      450:
      451:
      452:
      453:
      454:
      455:
      456:
      457:
      458:
      459:
      460:
      461:
      462:
      
```





PRIOR FILING DATE:	1958-02-09
PRIOR APPLICATION NUMBER:	60/075924
PRIOR FILING DATE:	1958-03-25
PRIOR APPLICATION NUMBER:	60/081045
PRIOR FILING DATE:	1958-04-08
PRIOR APPLICATION NUMBER:	60/035998
PRIOR FILING DATE:	1958-08-10
PRIOR APPLICATION NUMBER:	60/037000
PRIOR FILING DATE:	1958-08-18
PRIOR APPLICATION NUMBER:	60/035601
PRIOR FILING DATE:	1958-09-09
PRIOR APPLICATION NUMBER:	60/059803
PRIOR FILING DATE:	1958-09-10
PRIOR APPLICATION NUMBER:	60/059811
PRIOR FILING DATE:	1958-09-10
PRIOR APPLICATION NUMBER:	60/059812
PRIOR FILING DATE:	1958-09-10
PRIOR APPLICATION NUMBER:	60/100858
PRIOR FILING DATE:	1958-09-17
PRIOR APPLICATION NUMBER:	60/101922
PRIOR FILING DATE:	1958-09-24
PRIOR APPLICATION NUMBER:	60/106032
PRIOR FILING DATE:	1958-10-28
PRIOR APPLICATION NUMBER:	60/105304
PRIOR FILING DATE:	1958-11-20
PRIOR APPLICATION NUMBER:	60/125778
PRIOR FILING DATE:	1959-03-23
PRIOR APPLICATION NUMBER:	60/135635
PRIOR FILING DATE:	1959-06-15
PRIOR APPLICATION NUMBER:	60/145070
PRIOR FILING DATE:	1959-07-20
PRIOR APPLICATION NUMBER:	60/115658
PRIOR FILING DATE:	1959-07-26
PRIOR APPLICATION NUMBER:	60/149366
PRIOR FILING DATE:	1959-08-17
PRIOR APPLICATION NUMBER:	60/165455
PRIOR FILING DATE:	1959-12-07
PRIOR APPLICATION NUMBER:	08/918874
PRIOR FILING DATE:	1957-08-26
PRIOR APPLICATION NUMBER:	08/933821
PRIOR FILING DATE:	1957-09-15
PRIOR APPLICATION NUMBER:	08/965077
PRIOR FILING DATE:	1957-10-25
PRIOR APPLICATION NUMBER:	05/114844
PRIOR FILING DATE:	1958-07-14
PRIOR APPLICATION NUMBER:	05/136801
PRIOR FILING DATE:	1958-08-19
PRIOR APPLICATION NUMBER:	05/136804
PRIOR FILING DATE:	1958-08-19
PRIOR APPLICATION NUMBER:	05/136828
PRIOR FILING DATE:	1958-08-19
PRIOR APPLICATION NUMBER:	05/158342
PRIOR FILING DATE:	1958-09-21
PRIOR APPLICATION NUMBER:	05/180957
PRIOR FILING DATE:	1958-09-10
PRIOR APPLICATION NUMBER:	05/202088
PRIOR FILING DATE:	1958-12-08
PRIOR APPLICATION NUMBER:	05/253111
PRIOR FILING DATE:	1959-03-03
PRIOR APPLICATION NUMBER:	05/254460
PRIOR FILING DATE:	1959-03-09
PRIOR APPLICATION NUMBER:	05/254466
PRIOR FILING DATE:	1959-03-05
PRIOR APPLICATION NUMBER:	05/284666
PRIOR FILING DATE:	1959-04-15
PRIOR APPLICATION NUMBER:	05/332922
PRIOR FILING DATE:	1959-06-14
PRIOR APPLICATION NUMBER:	05/332923
PRIOR FILING DATE:	1959-06-14
PRIOR APPLICATION NUMBER:	05/333075
PRIOR FILING DATE:	1959-06-14
PRIOR APPLICATION NUMBER:	05/333077
PRIOR FILING DATE:	1959-06-14
PRIOR APPLICATION NUMBER:	05/333077

[illegible]

Alignment Score:	
Pair. No.:	2
Score:	20
Percent Similarity:	100%
Best Local Similarity:	207%
Query Match:	90%
Ds:	90%
Gaps:	3
Length:	32
Matches:	15
Conservative:	16
Mismatches:	30
Indels:	7
Gaps:	3

